IMPACT ASSESSMENT OF FINANCIAL MARKET DEVELOPMENT THROUGH THE LENS OF COMPLEXITY THEORY

JANUARY 2014

FSD Kenya
Financial Sector Deepening

Centre for Development Studies (CDS)
University of Bath

Dr Susan Johnson,
Centre for Development Studies, University of Bath, UK
and Dr Jean Boulton,
Visiting Senior Research Fellow, Department of Social and Policy Studies, University of Bath, UK
and Claremont Management Consultants Ltd

FSD Kenya
Financial Sector Deepening

The Kenya Financial Sector Deepening (FSD) programme was established in early 2005 to support the development of financial markets in Kenya as a means to stimulate wealth creation and reduce poverty. Working in partnership with the financial services industry, the programme's goal is to expand access to financial services among lower income households and smaller enterprises. It operates as an independent trust under the supervision of professional trustees, KPMG Kenya, with policy guidance from a Programme Investment Committee (PIC). Current funders include the UK's Department for International Development (DFID), the Swedish International Development Agency (SIDA), and the Bill and Melinda Gates Foundation.

Government of Kenya
UKaid from the British people
Embassy of Sweden Nairobi
Bill & Melinda Gates foundation

Table of Contents

Abbreviations

CBK	Central Bank of Kenya
CIS	Credit Information Sharing
CRB	Credit Reference Bureau
FSD	Financial Sector Deepening Trust, Kenya
IFC	International Finance Corporation
JTF	Joint Task Force (Credit Information Sharing)
KBA	Kenya Banker's Association
MFI	Microfinance Institution
PM	Project Manager
RCT	Randomised Control Trial
SACCO	Savings and Credit Cooperative
SG	Savings Group
SME	Small or Medium Enterprise

EXECUTIVE SUMMARY

FSD works to stimulate pro-poor change in the dynamic context of the financial market in Kenya. In doing so it engages with the complex, systemic, emergent and uncertain world that is the territory of complexity theory. This research set out to explore whether and in what ways evaluation and impact assessment can more effectively take into account this context.

Complexity theory emphasises that the world is:

- Systemic and synergistic.
- Shaped by history and context.
- Episodic in the way change happens.
- Evolving, where emergent features arise which could not have been known in advance.

Using this understanding of this complex world, the research was exploratory and had three purposes:

- To explore whether the contexts in which FSD works do indeed reflect this complexity.
- To enquire into the extent to which project teams responded to this complexity.
- To develop methods of impact assessment which assess impact in dynamic, evolving environments where there are many influences and actors.

The research focused on two projects – the development of a Credit Information Sharing (CIS) scheme for Kenya and the implementation of savings groups in West Kenya. It drew on and developed certain theory-based methods of impact assessment to examine impact both retrospectively and prospectively.

The first method traces the history of the project within its context through interviewing key stakeholders. The questions were designed to explore what happened, why, and what other wider factors were relevant. The respondents were also asked to consider what impact the project has made or is making. These core narratives were then triangulated, and, through second interviews with some respondents, any anomalies or areas of particular interest were further explored. Secondary data from written material was also used.

Using the resultant histories or traces of the project, we were first able to show that the project teams were responding to the complexity of the situation: dealing with unexpected changes in the wider context, with unexpected situations which had not been planned for, experimenting and customising approaches as necessary and seizing opportunities. So, the method both demonstrated that the world is complex in the way complexity theory infers and that project teams, in the main, are adept at responding to this complexity. This response to complexity is often achieved by project staff de-emphasising overly-constrained plans and processes that do not anticipate the importance of context and emergent change.

This method, which sought in an open-ended way to discover what did happen rather than to be constrained by what was intended to happen, also gave information about contribution to impact through following the sequence of events. This highlighted the impacts that, in the eyes of respondents, had occurred or were beginning to emerge. It produced indications of where corroborating information could be sought or monitored.

The second approach was more prospective in nature. Through interviewing groups of stakeholders, and exploring what was happening in relation to the project and its wider context, it was possible to collect 'narrative fragments' that gave evidence of emerging changes and trends that were still 'growing shoots' or 'weak signals' but were not necessarily yet amenable to quantitative analysis. This method was able to show the beginnings of impact pathways which went beyond the intended theory of change. It gave indications of where project teams might monitor some of these trends, where consideration needs to be given as to how these growing impacts might help or hinder the core intentions of the project and how this might inform programming practice as well as provide a more complex view of what had been achieved.

The report concludes by discussing the implications of these findings for programming, evaluation and impact assessment. It suggests ways in which FSD can embed these practices in the management processes of the projects and programme as a whole. This would allow FSD both to undertake effective impact monitoring which better informs project and programme development and to provide input into retrospective impact assessment studies.

Chapter 1

WHY COMPLEXITY? THE COMPLEX NATURE OF FSD'S WORK AND CONTEXT

The Financial Sector Deepening Trust was established in 2005 to support the development of financial markets in Kenya as a means to stimulate wealth creation and reduce poverty[1]. The financial market is undergoing rapid and far-reaching changes and there are a range of inter-related driving forces and uncertainties which will shape how the market evolves, as FSD's strategy clearly demonstrates. These include: growing communication networks offering new channels and opening the market to new entrants; changing demographics with a growing youth sector; urban migration; decentralisation; changing policy and regulatory frameworks.

FSD's approach is to influence the evolution of the market in ways that are pro-poor. In order to do this FSD engages in a range of projects and activities. FSD's work - in being innovative and working in dynamic contexts to stimulate change - engages with the complex, systemic, emergent, uncertain world that is the territory of complexity theory. FSD exemplifies many examples of response to this complexity. It acts to analyse, anticipate and 'foresight' [2] the future – making judgments as to how the market and its determining factors will develop. It builds relationships and engages with key stakeholders in order to gain insight and to be in a position to influence future strategies. It seeks to be adept at seizing opportunities to catalyse changes that support poverty reduction. It supports innovation and invests in and demonstrates new ways of working. And sometimes it persists with and supports ventures that are less popular with the mainstream but nevertheless are judged to be important for poverty reduction (see Box 7).

FSD faces a number of challenges in evaluating and assessing the impact of its work as a whole, as well as for particular projects and programmes. The focus of the research described in this report is to investigate what complexity theory has to offer to considerations of how FSD can evaluate and assess the impact of its work. However, this inevitably raises issues related to how programmes are designed and the report therefore also addresses aspects of design and implementation.

This research addresses the conjuncture of complexity thinking with programming and impact assessment in a development context. Complexity thinking has been of increasing interest to development practitioners in the last few years. However, there is as yet little in the way of commonly-accepted approaches in which it is applied practically. The arena of impact assessment currently attracts hot debate about approaches and methods. Bringing programming and impact assessment together to investigate their implications for FSD's work is an innovative step; with no pre-existing methodologies to apply, the work is exploratory. Indeed it was first important to consider whether complexity thinking was in fact relevant to FSD before turning our attention to methods and frameworks. Our intention was to draw out key features of complexity thinking which are of value to programming and impact assessment and propose ways forward - a "complexity-informed" approach.

The report proceeds as follows: section two gives an overview of complexity theory as a world-view and explains the questions this raises for this research. There follows a discussion of methodological considerations with complexity in mind.

The third section describes the case studies and demonstrates in what ways they accord with complexity thinking and how they inform good practice in complex contexts. The fourth section then considers the implications of complexity thinking for operations, focusing on evaluation and impact assessment in some depth. The fifth section, taking this into account, concludes with a proposal for how FSD can improve its monitoring and evaluation work in complex contexts.

The report provides the reader with the core argument. The Annexes offer more detailed background on the state of the debate on impact assessment and its relation to complexity theory. They also give greater detail on background work for one of the case studies and more detail on the proposed approaches.

---

1 FSD Kenya Strategy 2011-2015.

2 Foresighting is a recognised strategic technique that supports judgements about future markets and their determining factors. See Clayton, A., W. Wehrmeyer, et al. (2003). Foresighting for Development, London: Earthscan Publications.

Chapter 2

COMPLEXITY THEORY AND ITS RELEVANCE TO FSD

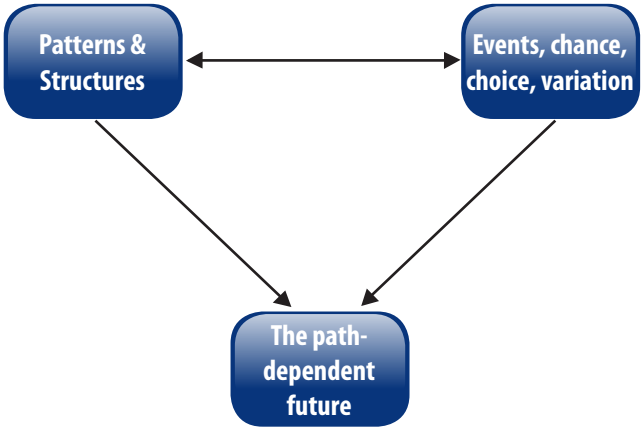
2.1 COMPLEXITY THEORY: PRIMARILY A WORLD-VIEW

Complexity theory provides a world-view, derived from the science of open systems and of non-linear interactions between elements at the micro-level. It describes and articulates the characteristics of a world that is interdependent and shaped by history and context, a world that is dynamic and subject to rapid shifts and the associated emergence of new features. It provides a fundamental, scientifically-informed 'ontology', that is to say, a description of the way the world is - a perspective on 'being'. It also considers the processes of interaction and change – a perspective on 'becoming'.

Complexity theory provides explanations of the way in which patterns and order can emerge. It brings together the physical sciences and the theory of evolution. Non-linear systems thinking in its various forms (von Bertalanffy, 1969) points to the essential interconnectedness of the social and natural world. It emphasises that few outcomes have single causes and few interventions or events have single outcomes. More than that, interconnections cannot in general be treated as independent of each other, as summative. Many events and underlying embedded cultural and economic 'norms' work in synergy with each other; that is to say that their impact is more than – or less than - the sum of what their impacts would have been if each factor had operated in isolation.

Complexity theory, however, goes beyond systems thinking, which tends in the main to focus on situations, contexts and interventions as if they were stable (P. Allen & Boulton, 2011). Systems methods in general provide a description of the present; complexity theory in addition pays attention to the *dynamic* and *evolutionary* nature of the context, and the way new features can emerge. Many factors can contribute to the emergence of new features: specific events, local variations, shifting environmental factors, and reflexively-developing intentions and strategies. Complexity can be considered as a dynamic perspective on the way institutions are created and modifying or destroyed by events and the particularities of people, decisions and chance variations.

Figure 1: interactions in a complex world



As illustrated by the diagram above, complexity theory emphasises the interplay between currently- instituted **patterns** of **relationships** and **events**. Patterns exist at differing, although interdependent levels in society and combine social, political, economic and cultural aspects. There are also the instituted patterns within organisations. Events and variations which destabilise such patterns can include the arrival of a new local leaders, a successful business idea emerging from a local entrepreneur, war, drought, new technology, and influences from the media. This idea of the dynamic relationship between patterns and events, between 'science and history' (P. M. Allen, 1997) as it is sometimes described, is central to complexity thinking and changes the way we think about the way things are and the way they come to be.

Complexity thinking stands in contrast to the idea of the 'world as a machine' which is the world-view (ontology) underpinning many conventional theories of change, strategy development and evaluation. The idea of the world as predictable and measurable, where causal chains can be identified and form the basis of plans and evaluations, is based on the mechanical science of closed systems with linear forces the relevance of which is brought into question in complex, uncertain, interconnected and fast-changing contexts such as those familiar to FSD.

Box 1: The complexity worldview

The complex world is:

- Systemic and synergistic
  - It cannot in general be treated as a machine, deconstructed into separate parts connected by independent linear forces
  - Non-linear interactions lead both to synergistic and antagonistic interactions
- Shaped by history; path dependent
  - History matters; the past shapes the present and the sequence of events – the path - is significant to what happens next
- Sensitive to context
  - The *particularity* of contextual factors and the particular detail of events materially impacts outcomes and cannot be ignored
- Episodic
  - Change is episodic not smooth or incremental. There are periods of apparent stasis and times (tipping points) where radical change occurs
- Emergent
  - New features can emerge at tipping points (and others disappear) and such qualitative change cannot be predicted
  - 'Things' are dynamic, not static and the future cannot be known in advance; there can be unintended outcomes and unexpected shifts and events

Box 1 summarises the features of a complexity world-view. The underlying precept is that interactions cannot be treated as linear and that linearity is a special case seen only in stable situations where there is little potential for change. Non-linear interactions at the micro-level are what lead to the features of a complex world – to synergies and to the inability to treat interactions and effects as if they were independent of each other. Non-linear interactions lead to episodic change and to the growth or decline of currently-established patterns at the meso- and macro-level.

Episodic change means that change advances through periods where there is little to show, but where interactions and factors may slowly be changing Then there are times, so-called 'tipping points', when change can be rapid and radical. In periods of seeming stasis, it can be important to persist in moving towards goals and intentions, to 'build the ground' towards change even when there is little evidence of change. Equally, it is important to react effectively to the opportunities and threats at critical junctures (points which seem to have the potential for 'tipping'). This is where change can be radical and new features and indeed eras can emerge. In other words, it is easy to give preference to the importance of tipping points[3] in reviewing change and downplay the less-obvious role of persistence at times where there seems to be little evidence that anything is happening.

These features of a complex world mean that, in order to understand how things happen it is necessary to try to follow the path, the time-sequence of events, and to assume that this sequence is unique to the circumstances. It is only by tracing this unique path that it is possible both to understand *why* something had impact and, *indeed*, to have confidence that it did indeed contribute to impact. This thinking underpins our approach to retrospective impact assessment with complexity in mind in which we develop a method akin to process tracing, as will be discussed below.

Furthermore, in looking forwards, complexity thinking emphasises that change happens because stable balances between factors are disturbed so that effects not seemingly important start to grow and develop. The classic example of this thinking is the ecological balance that existed during the time of the dinosaurs. Small mammals seemed insignificant to the ecology, hardly

Box 2 : An example: pro-poor financial market development in Kenya – identifying and understanding a "tipping point"

In the mid-2000s there was a qualitative change in the banking sector's understanding of its market. Whereas bankers had previously been seeking to compete over those people they thought bankable, they now recognised that there was a huge market of unbanked people whom they could seek to serve.

A tipping point is a point at which new patterns emerge, where previous patterns are broken and new patterns come into place. At least two factors are identifiable as having contributed to this change. First, Equity Bank changed its fee-charging approach away from monthly ledger fees to charging per transaction This was mainly based on withdrawals, while also allowing zero minimum balances. This was a key moment in making ordinary bank accounts affordable to the majority and it attracted many more customers as a result. It demonstrated to the rest of the banking sector that there was a large market out there which it had not to date considered bankable. Second, the publication of the FinAccess survey results in 2006 also made an impact on the banks' understanding of their market. The survey revealed that only some 22% of adults had a bank account, leaving a potentially huge market to play for.

However, to arrive at such a tipping point it is necessary to explore the other causes and conditions that contributed to it. What 'built the ground'? How did factors converge in ways that produced synergies? Equity had been developing its business model and undertaking a lot of product innovation and development, taking the feedback from its customers seriously. Support for this process was also due to FSD Kenya's involvement via MicroSave. It was Equity's identification of itself in the late 1990s with the idea of serving this microfinance market that propelled its vision forwards. But Equity's engagement with the low-end market would not have happened had it not identified itself with this sector serving low-income and unbanked people in contrast to dominant commercial banks at the time. Across the globe, lessons in microfinance were also being learned as a result of the experiences of – those Kenyan microfinance institutions such as K-REP, KWFT, Faulu who were some of the early experimenters in lending to poor people.

While the Kenyan microfinance sector that was embodied in a group-based, credit-led model had seen rather limited success in terms of market outreach, it nevertheless contributed to the development of the conditions from which Equity and its approach later emerged.

In addition, there were other factors that contributed to change. For example, the role of particular leaders and entrepreneurs, such as Equity's James Mwangi, as well as the political economy of Kenya which also shifted significantly in 2002 when a new president was elected.

Based on (Stone, Johnson, & Hayes, 2010).

[3] A note on language – we define tipping point as a place where there are qualitative differences between the before and after state – new qualities have emerged. We use critical juncture to signify a point at which things may tip – but, depending upon actions and other factors, may not.

important at all. But when the asteroid crash triggered changes in the climate changes that led to the extinction of the dinosaurs, there was room for small mammals to flourish. The point is that in order to explore emerging impact pathways it is necessary to be sensitive to 'weak signals' – that is to look for signs of change before they become established into new patterns. This underpins the approach presented here to looking prospectively at emerging impact pathways as is discussed below.

2.2 METHODOLOGY IN A COMPLEX WORLD

A central goal of this research is to experiment with and develop methodologies which allow programming, managing, evaluating and assessing impact in a complex world. In section 4 we describe these in-some-cases experimental and developing methodologies. In this section are summarised some of the over-arching features that such methodologies must reflect in their design.

Limits to universality in methodology

One of the key understandings of the complex world is that situations are materially affected by context and by history. So complexity theory does not lead - and indeed *cannot* lead - to universally-applicable frameworks of implementation or analysis. Indeed any framework that claims to be universally-applicable should be viewed with caution. Rather, complexity thinking leads to the definition of certain *principles* and contingent guidelines – leads to *meta*- theories of change and *over-arching* considerations for evaluation and impact assessment – rather than to detailed micro-theories that have generic, more universal application.

Limits to universality in criteria for success

The lack of universality has consequences for trying to establish criteria for success. These criteria will be determined by the particular conditions (social, political, organisational, institutional etc.) and are thus not easily generic. They also cannot capture a complete picture, especially since they will be applied at a point in time and new factors may emerge during the course of the project or programme. Indeed something that is regarded as a 'failure' at a particular point in time may then turn into a 'success' at another point in time -or vice versa.[4]

A need to trace the pathways of events and impacts and take note of the dimension of time

Complexity theory emphasises that the way in which change arises is contingent on the detailed processes, different in each case. So, in order to understand this heterogeneity of outcomes and results across space, time, social context and so on, as is often the requirement for impact studies, it is necessary to understand the 'how' as well as the 'what'. Complexity-informed empirical methodologies must therefore trace the pathway, the history of how change arises in the particular case.

A need to have ways to spot 'weak signals' and signs of emerging change

Complexity theory sees change as arising at the microscopic level, in the detailed interactions in particular circumstances. Change may then aggregate from these beginnings and become more sustained. However, to understand change and spot emerging impacts which might inform programme changes or decisions to monitor new features, there needs to be ways of spotting signs of such emerging change.

How precisely to undertake intervention to create change or assess impact will in practice differ between contexts and for different types of programme. This exploratory research work has therefore led to suggested guidelines and considerations for processes to be piloted. It outlines further explorations to be considered and these will be described in overview in section 4 and in more detail in the annexes.

[4] An example in the microfinance field in Kenya is either (i) Equity Building Society whose failure and virtual insolvency in itself set in place the conditions under which it was necessary for it to transform itself; or (ii) along similar lines the withdrawal of external funding from Partnership for Productivity in the 1980s which resulted in the managed ASCA model evolving in Central Province and developing a model with outreach to over 20,000 (Mule, Johnson, Hickson, & Mwangi, 2002). Or the example of the Western financial crisis in 2008, that a booming banking sector– given time proved to be a source of failure for the sector as a whole – indeed part of a long history of cycles in this sector.

Chapter 3

THE CASE STUDIES

3.1 THE PURPOSE AND METHODOLOGY OF THE RESEARCH

In order to trace the details of change processes consistent with the application of complexity principles, the research work for FSD, based on two cases studies, was designed to explore the history, context and pathways of the projects and hence to chart the sequence of events involved in implementation and the range of factors contributing to the way this occurred in practice. It sought evidence of outcomes and evidence of emerging change and impact. The purpose was:

- To investigate whether the description of a complex world as defined by complexity theory (set out in Box 1) fitted the context and experience of the projects in the way implementation took place and change was (or was not) happening.
- To consider how the implementation processes the projects engaged in actually responded to the complexity of the context (as set out in Box 6).
- To examine what could be learned that was relevant to project design, evaluation and impact assessment.

The two projects chosen from FSD's portfolio are very different in both their character and form and have different final impact goals of growth and poverty reduction. They therefore capture some of the diversity of FSD's portfolio.

The Credit Information Sharing project commenced in 2008 when FSD recognised there was a gap in the capacity of the sector to implement the Central Bank of Kenya's (CBK) newly-mandatory negative credit information sharing guidelines. It funded a project manager based in the Kenya Banker's Association to take forward implementation through a Joint Task Force. This developed into a second phase from 2011 which aimed to develop a full-file sharing system - that is complete credit records which capture positive information about client loan repayment as well as negative information about bad debts.

The Savings Group project in which FSD has worked with CARE since 2008, aims to provide small-scale savings and loan services to poor people and to those excluded from formal financial services. It does so through training groups of approximately 30 people in a methodology for saving and lending. Groups are trained for a year, which takes them through one cycle of saving, lending to each other and then undertaking an action audit through the sharing out of the fund, and then starting the cycle again if they so wish. FSD's involvement in the project has experimented with alternative delivery channels for the training in order to test ways of lowering cost and understanding the impact of this on the quality of group formation and sustainability. Finding approaches which significantly lower the cost per member is intended to allow for scaling up of the dissemination of the methodology and achievement of mass inclusion in this basic financial service. The second phase of the project since 2011 is intended to lead to decisions on what approaches to adopt by the end of 2013. In the table below, we identify the differing characteristics of these projects, and in the next section, we describe the results of our research.

Table 1: Comparison of case study project characteristics

Credit information sharing (CIS)	Savings groups
▪ **Impact pathway** expected to lead directly to lower interest rates for SME finance as banks information costs are lowered; they implement risk based pricing, and information rather than using physical collateral. This leads to SME development and growth.	▪ **Impact pathway** expected to lead to direct impacts on access to appropriate financial services for poor people and thereby to poverty reduction. This is achieved through use of service to increase incomes and reduce vulnerability as a result of consumption smoothing.
▪ **Project design** much more open and responsive to proposed outcome which is full-file credit information sharing.	▪ **Project design** much tighter as to how the outcome – i.e. appropriate financial services delivered through savings groups – is to be achieved and testing of effectiveness via different delivery channels.
▪ **Small 'n'** – only one initiative	▪ **Large 'n'** – i.e. large number of instances of the intervention at work.
▪ **Prior experience of implementation very limited** – experience from other countries drawn upon via consultants but no direct experience in Kenya.	▪ **Prior experience of implementation** elsewhere where earlier learning and adaptation can be built upon.
▪ Evaluation and impact assessment: project completion report of phase 1 available but project continuing – no time lapse to see what has emerged/ resulted from project over time. ▪ **Wider impacts** are also relevant (e.g. stability of financial sector due to fewer 'bad debts', increased trust between banks etc.)	▪ **Evaluation and impact assessment**: a range of studies have been conducted. ▪ Time: a significant time has elapsed following earlier phases – so time for wider (and narrower) impacts to emerge and be visible. ▪ Wider impacts: the potential for wider impacts to occur – on the socio-economic and financial services landscape - over and above the intended impact.

METHODOLOGY

The focus on the research was to develop a history of the project and of what led up to the project – to allow investigation of the detailed pathways that have unfolded. This is a version of process tracing as we will discuss later. In addition we asked interviewees about any evidence of change that was beginning to emerge so that we could identify any prospective 'growing shoots' of change, that is, evidence of new impact pathways.

We not only looked 'inside' the project, we also explored the wider contextual and historical factors which could or have shaped and influenced the project design, and contribute to or constrain potential success. We explored the **past** – what was happening in the context before the project started, that is what earlier factors were relevant, be they cultural or due to other interventions or environmental or political factors. We asked questions about the wider context in the **present** – other players, changes to social, political, environmental and economic factors. And we looked forwards to the **future** – key events such as elections or private sector investment, other planned changes and interventions, future potential opportunities and threats.

The research involved primary interviews with project implementers and stakeholders, and review of secondary project documentation. The objective was to glean the key events and key determinants of the project path through triangulating these accounts. In some cases second interviews were undertaken to check information, or explore certain points in more detail or to discuss issues where there were differences of opinion. In the case of the savings groups, interviews were with groups of stakeholders as well as with individuals.

The research was undertaken over a total of eight days of field work in Nairobi and in the field. This was not therefore intended to be a detailed impact assessment of either project, but was designed to investigate how change was occurring; how this could be better understood using the insights of complexity theory. An additional aim was to propose an approach for FSD as a programme that could better capture this complex change.

In the following sections the experience of the projects is described. The inserted boxes give a commentary on these events from the perspective of complexity theory.

3.2 CREDIT INFORMATION SHARING

3.2.1 Background

The project works primarily with key stakeholders – financial institutions, the Central Bank, and credit bureaux, and engages with other stakeholders as need arises. It is strategic in nature, with a clear long-term goal to establish a credit-information-sharing system which contains positive credit information about borrowers as well as information on bad debtors as initially mandated.

Since it is obviously a unique, one-off project in Kenya it has to manage its way forward, responding to the emerging needs that implementing an information-sharing system virtually from scratch in the local context has required. It has had to manoeuvre to address the challenges this raises, while at the same time seeking to move opinion towards the benefits of sharing positive as well as negative information.

3.2.2 The story of the project and its back-history

FSD started its first phase of involvement in 2008 to bridge a gap in the implementation of the newly- mandated negative credit information sharing guidelines. FSD funded both the project manager based in the Kenya Banker's Association (KBA) and related activities to take forward implementation including through a Joint Task Force. This grew into a second phase project in 2011, intended to develop the system further and bring about full-file information sharing.

In this section, the story of the development of credit information sharing in Kenya is summarised, taking as a starting point the issue of 'political loans' in the 1990s and ending with the passing of legislation for full-file share at the end of 2012. Not every feature is included; the narrative is designed to pull out what are seen to be key moments and key contributory factors in moving forwards. We provide a commentary through a 'complexity lens'.

The early 1990s: The banking sector had a history of lending in the 1980s and 1990s which resulted in three phases of bank failure. This was as a result of the political ownership of some banks and political basis of many lending decisions. For the banking sector the difficulties of managing bad debts and ensuring that debtors did not simply move between banks was of prime concern. It was also very politically sensitive. Banks had become very conservative in their lending requiring 100% collateral. At this time a number of stakeholders had different objectives but common aims which were complementary - but MPs were pulling against this. These stakeholders included:

- The CBK, which was concerned with the stability of the banking system due to the impact of non-performing loans on bank collapses. The bank saw the potential for a credit reference scheme to stimulate growth through supporting increased lending.
- The KBA as an industry association was concerned to find ways to reduce the bad debt portfolio.
- Credit bureaux: one bureau – Metropole (1996) - acted as an agent for Dunn and Bradstreet, developing credit reports on Kenyan companies wanting to trade abroad. Another bureau – Credit Reference Kenya – started operating in the early 1990s as a debt collector for banks.
- The IFC, which had been working for some time supporting the development of CRBs in developing countries and were active in Kenya.

These motivations pulled together **synergistically**, apart from MPs who were resistant

These motivations all pulled together **synergistically** but MPs were resisting the issue. They were concerned that the finger would be pointed at them if their bad debts and their bouncing of cheques presented during high profile *harambees* (fundraising events) came to light.

The **history** of 'political lending' and the political and cultural importance of *harambees* meant that MPs were resistant to credit references in case they were discredited.

Slowly creeping forwards 1992-2006: in 1998 four banks collapsed. This gave the impetus for CBK to establish a credit reference process for very bad debts which was allowed by common law. By 2000, six banks were taking part. The debate developed as to how to establish a more fully-established credit reference system. The CBK did not wish to host it, and the banks wanted CBK to license a CRB. However the non-performing debt crisis eased off in the late 1990s so the heat was off for establishing a CRB. IFC suggested to the two credit reference bureaux, CRB Africa and Metropole, that they should establish a CRB association to use as a lobby group. They approached KBA around 2002 to push for the sharing of information.

The impetus for a credit reference scheme waxed and waned in relation to the level of bad debts – **path dependency**

In 2004 the Banking Act was amended to *allow* sharing of negative information, but did not make it mandatory. By 2006 Barclays in particular was making some unsecured loans as a result of informal information sharing.

There was **persistence** towards the goal, driven in part by the motivations and lobbying of the credit bureaux – and a final breakthrough in 2004

A key decision by parliament was made in 2007 for mandatory sharing of negative information. One of our informants reported that in May 2007, the Leader of the Finance Committee of Parliament asked for a meeting with him as MD of Metropole to have him explain why sharing of negative information was so important – its relation to controlling interest rates and impact on poverty if loans are less driven by collateral. He also wanted to know that it was not aimed at embarrassing MPs. This contributed to the passing of the bill.

A tipping point

The regulations then had to be written and a Joint Task Force (JTF) organised with IFC playing a key part with KBA and CBK was formed in November 2007 – but little progress was made.

This was a **critical juncture** – a point where there was a sense of frustration that things were not moving forwards – FSD **seized this opportunity.** There were several reasons, coming together, why FSD was selected and why the time was right.

Arrival of FSD in late 2008. At a three-day workshop focused on implementation it was clear that more focus and resources were needed. FSD offered to fund/help. Their motivation was distinct from others' in that their focus was the potential to increase lending to SMEs.

FSD's offer to support implementation was accepted as the key IFC person was also withdrawing at that time, and because stakeholders saw how much there was still to implement. Also FSD was known and trusted. However, the initial recruitment of a project manager from the private sector failed and FSD approached the current PM who had been working at the CBK on this issue and had involved FSD in the workshop.

The Project Manager (PM) started work in August 2009 seeking to regenerate momentum. Although the banks were supposed to be compliant with the guidelines by February 2009, this had not been achieved. His prior connections to stakeholders in CBK and KBA helped generate influence and move things forward. Together with his reflective and anticipatory approach, he was well-placed to keep attention on the longer-term goal of sharing positive information, to anticipate what was required and to plan ahead.

The choice of PM was **critical** due to his forward-looking skills, his persistence towards the goal, his ability both to seize opportunities to influence and plan critical moments ahead (e.g. conference) and his networks.

The project had three pillars: communication, specifying protocol for data-sharing and sorting out legal implications. The idea of having a champion in each bank was used to provide focal points and move things forward. However, there was a strong negative public perception that this was a 'black-list' which was not helped by the way the press reported. Initially while the law mandated sharing of information it did not require that banks used it so few banks were buying credit reports. The first CRB was licensed in February 2010.

This was a systematic approach

The project had to **adapt** to this negative issue of black-listing and endeavour to establish a campaign of influencing and turning this negative publicity to its advantage.

A pilot exercise in reporting to the CRB in Spring 2010, was *'particularly depressing'* in the eyes of FSD. Large numbers of records were rejected by the CRB and banks were still nervous about sharing the information. The PM *'played the regulator card'* and emphasised sharing was mandatory.

Another example of **adapting** in light of a negative response.

The second CRB, Metropole, was licensed in April 2011.

A planned-for, carefully-considered key event: East African Regional Conference July 2011. The project brought World Bank and Bank for International Settlements experts and others from all over the world, along with key stakeholders in Kenya and the rest of the region. This showcased Kenya's progress but was also used to show the importance of sharing positive information as internationally-recognised best practice. The event built support for this through sharing examples of how this had been effective in other countries. This was widely reported as a really critical moment – perhaps the most important turning point – by all interviewees. It launched the national task force and discussions on how to design the next phase of the project.

This was seen by many as a very important **critical juncture** for the project – which was **anticipated**, designed and worked towards on many fronts – seizing opportunities and working intentionally.

A critical moment came in July 2012 with the mandatory sharing of full-file information. The CBK was keen to head off the threat of an interest rate cap that was being discussed by Parliament - a recurring issue given that interest rate spreads had remained high. The CBK met with MPs and one in particular was interested to understand how full-file sharing would contribute to lower spreads by reducing costs for the banks. The amendment was passed the same evening although the MP had seized the opportunity to have it included in the Central Bank Act rather than the Banking Act.

This leapfrogged an initiative of the JTF to bring the largest banks together to develop a code of conduct for voluntarily sharing full-file information. The legal amendments had also been made for MFIs to share negative information but not positive. Together, these developments mean that there is a lot of work to be done in harmonising the legal framework and interpreting the regulations appropriately and in building the cooperation between the banks required to make full-file really effective.

This was a curious critical juncture. It came almost too soon and cut across the gradual building towards this point that had been happening. It had to be adapted to.

3.2.3 Was the context complex?

This history and background to the project demonstrates a number of features of complexity theory. We follow the structure given in Box 1.

First, a number of factors came together **systemically** to create the conditions for the move towards a credit sharing scheme. Differing stakeholders wanted credit information sharing for differing reasons, but these reasons did not pull against each other. The Central Bank wanted to ensure the stability of the banking system; the banks wanted to improve the quality of their lending, and the credit bureaux were driven in part by the potential for extending their operations. Stakeholders such as FSD wanted to improve the potential for SMEs to get credit at lower interest rates through information rather than physical collateral; and politicians – who held out for some time – eventually saw this as a way to address the need for lower interest rates.

Second, the **history** and **context** and sequence of events were important. The banking crisis in the early 1990s led to an interest in credit referencing. Then in the late 1990s, the heat went off, slowing things down. The particular issue of political loans being used to 'lock in' political support together with the importance of harambees meant that politicians were not entirely positive about credit sharing. How and when credit sharing was agreed and implemented was therefore constrained by the particular history and culture of Kenya. Another country would have had a different sequence of events and different conditions to address and different moments to seize when conditions became favourable to take action. The story of particular relationships – part of the specificity of the context – was also important: how influences for change and involvement of particular people (such as FSD's involvement and the subsequent selection of the project manager) came through particular conversations held at particular times or through particular events.

Third, the history demonstrates the **episodic** nature of change: the way change in attitudes often happened imperceptibly over several years and then key events (which had been built towards by many stakeholders over extended periods of time) moved forwards suddenly – with the 2004 bill, the 2008 bill, and the sudden agreement of positive sharing. There were a number of **critical junctures** in the story – some, where there was a strong sense that, if these moments were seized, the project could make leaps forwards. Some of these were constructed and planned for (such as the international conference); others came as a result of influencing and slow change, but came unexpectedly – for example the final legislation for sharing full-file.

Finally, the history shows how, after a critical juncture, new features can **emerge**. For example, credit reference reports have now come to be used in

the selection of people for public office – this is a new emergent feature which was neither expected nor intended.

3.2.4 Did the project design and project team respond to the complexity of the situation?

The project team, and indeed the wider stakeholder group demonstrated in many ways, both in their actions and in the way they related the history of the project in interviews, that they were attuned to the complex nature of the context.

We set out the behaviours required to deal with complexity in Box 6 in section 4. The episodic and emergent nature of change requires adaptation and the ability to seize opportunities and respond to unexpected events and consequences. Equally, there is a need to set long-term goals and aspirations and persist in moving towards them – otherwise implementation becomes blown by the wind – adaptive and responsive yet unfocused and non-strategic.

First, the team was very clear about the ultimate goal to which it **aspired** of achieving full- file information sharing and persisted in this goal, operating purposefully to "build the ground" to this end even while they were still in the midst of getting the negative information sharing system operational.

The Regional Conference in particular was a strategic and **anticipatory** and very important bid to put this goal before the sector and engage with it. This was a planned and intentional way to try to create a 'tipping point' and was deemed to be successful in this respect; interviewees reported that this was the point at which the wider stakeholder group really started to believe that full-file sharing was both desirable and possible. The conference created significant legitimacy for this goal while also putting pressure towards the effective implementation of the negative system and creating the space for a new, wider stakeholder group that could take this goal beyond the banks alone into the other financial institutions and credit providers.

The team also demonstrated how it could **adapt** to the fact that the legislation for full-file sharing came somewhat unexpectedly and cut across the development of a platform for voluntary implementation which was almost in place when the mandating of full-file was agreed by parliament. This in turn has created new challenges to be managed as the legislation is somewhat piecemeal and not at all comprehensive.

The project, being one of a kind, is entirely **customised** to the particularities of the Kenyan context while drawing on learning and expertise from elsewhere.

3.2.5 What the case study demonstrated about impact

In particular in the second round of interviews, we were keen to explore what impact the interviewees felt that the project had achieved or was achieving and what impact information was being collected and could be collected. We were also interested to see – methodologically - how effective the interview processes were in gleaning information about emerging impact.

As is apparent, the negative information sharing is in the very early stages of operational use and the banks are only now learning to operate both to supply and use the information on bad debtors. To date, this information has been used by some banks as a means to encourage people to repay their loans since debtors are informed that their case will be reported if they do not pay. It is also used to screen if people have outstanding debts in other banks – in which case they will not be granted a loan. It is also used by some banks at the appraisal stage of loans to back up decisions. However, the very basic way the information is being used to date has demonstrated the lack of capacity within the banks to use the more detailed information that will be available in full file reports to deliver risk-based pricing.

There is some indication that the system has improved loan repayment performance although bad debts were in fact rising in early 2013 as a result of other economic factors. This in turn demonstrates the difficulties of assessing impact in this area by focussing solely on the prevalence of bad debts[5] and interest rates. However, it demonstrates that the detailed processes through which these changes have the potential to affect interest rates need to be carefully traced if the impact of its contribution is to be appropriately assessed in future.

At present monitoring is at the level of use of the system: for example, the number of reports being requested from the credit bureaux. Establishing impact requires working along the impact pathway through the bank's systems to see how these are being used/not used. It also establishes how they are affecting (or not) the ways in which both repayment enforcement and lending decisions are being made, along with monitoring the numbers of loans involved and to whom they are being made.

An unexpected use of credit reports has been their use in screening applicants for political and public office, which has arisen as a result of the integrity clause in the new Constitution. This has also been adopted for new directors by the banks themselves, and appears to be making its way into other parts of the private sector. This is an unexpected impact in terms of timing which also has positive feedback effects on the public credibility of the CIS system.

In terms of evaluating the research process, asking these questions about impact during the interviews was useful in several ways: it brought attention

[5] The PM also indicated a concern that loans may be being paid by resorting to borrowing from the informal sector and that this might be an issue to monitor. This is a useful example from a complexity perspective. It highlights an unexpected impact pathway that could be set off by the system. An appropriate approach would be to bear this in mind and note down any anecdotal evidence to suggest that it might in fact be happening, and then if this started to appear to be important enough, set about ways to monitor it more systematically.

to what information is known (requests for reports), what is emerging and could be monitored (increase in loans to SMEs, increase in informal loans). It also showed what requires a forward-looking process tracing approach to monitor and see if there are any indications of contribution (project events correlated with interest rate movements).

3.3 SAVINGS GROUPS

3.3.1 Background

The first phase of FSD's support to CARE's Savings Group project ran from 2008 to 2011 and reached 125,000 people in 4,500 groups. Beyond direct implementation, the goal of the project was to test different delivery channels for training groups in the methodology in order to lower the costs of delivery and establish effective approaches for a national scale up. The project is now in a further phase which is further testing delivery systems and consolidating learning (including moving to a more challenging area such as Marsabit). It had reached a further 80,000 people in some 3,000 groups by mid–2012.

3.3.2 History of the project

The history of the way in which CARE staff started implementing SGs in Western Kenya in the early 2000s is akin to the story of many types of initiatives which take an idea, try it out seek to adapt it and assess what will work in the local context. CARE had worked with mainstream microfinance models with little success and since it was also operating with programmes in agriculture and health, the organisation wanted to find an approach to financial inclusion that could work with poor and rural communities.

CARE started working with savings groups in Kenya in 2003, funded by USAID. The idea of savings groups was new to Kenya but had been implemented by CARE in Zimbabwe. CARE started implementing in Rusinga Island, but field staff were 'laughed at' by the local communities who were already receiving a lot of aid of various kinds. They mocked the idea that saving their own money could be of benefit.

This response was **path-dependent**, emanated from the 'aid culture' that had been established

The CARE team moved on to another area, Suba. Although there was less current development work in Suba, this was a region in which SACCOs had collapsed and people had lost a lot of money. So, in the fishing communities targeted, there was little trust and a real concern that any money would be lost. The CARE team gained their trust by reducing the savings collection interval from one month (as set out in the Zimbabwean manual) to one week and finally to one day.

The team **adapted** to the difficulties in Rusinga Island, by moving elsewhere and then had to **customise** the approach to deal with the (**path-dependent**) results of SACCO collapse

Over time, other adaptations were made. As demand grew (and Rusinga Island now demanded their own savings groups), the project team started to involve local trainers who were not paid but were given bicycles to help them travel to the groups (and which they were also allowed to keep). The training manual was adapted for their use.

The team **experimented** with implementation and adapted and **customised** the approach to training

There was tension between the institutional view (within CARE and within the wider academic community) that there is a 'right' way to implement savings groups (based on methods developed in Zimbabwe) and the need as felt by the project team to customise the approach for the communities they were working in. This tension was so strong that at one point the project was nearly cancelled as it was felt the project team was 'doing it wrong'. The situation was only resolved with the intervention of an expert in savings groups (Hugh Allen) who visited the project and validated their approach (he said it was the most innovative he had come across).

This tension highlights the difference in mind-set between adopting a machine view (right way, minimum experimentation and customisation) and an approach congruent with complexity (adaptation, customisation, experimentation)

The story highlights the importance of persistence – both within the communities which had to find a savings schedule that worked for the fishing communities), and within CARE, which had to battle with the feeling that the project was not implementing savings groups in the 'right way'.

The **persistence** of the project team to implement the process in ways they felt it would work – taking on CARE management at times.

Involvement of FSD

The CARE PM saw an advertisement for a call for proposals from FSD to which she responded. Discussion with FSD took over a year to arrive at a final proposal. FSD's interests were to find a cost-effective methodology to scaling up SG implementation. This in turn involved innovative thinking which applied elements of thinking about markets. It considered which actors might be available to implement including private sector businesses; and methods of structuring incentives that would support effective delivery.

Building on her prior experience, the PM was challenged to produce an approach to scale-up where different delivery channels could be compared (local trainers managed by franchisees or faith-based organisations and direct support from CARE).

There would seem to be a congruence between FSD's focus on comparing pilots and CARE's view of the need to customise and experiment.

While the initiation of the project therefore took considerable time, FSD offered a focus on experimentation but with a clearer research focus so that learning could be identified. This evidence-based approach to experimentation has been important for the project and is deemed to have created a research-based approach within the CARE team. This careful approach to considerations of scale-up and efficacy has supported the growth of savings groups and helped those involved to learn which methods are most effective.

FSD brings a more careful, research-based approach to experimentation which has changed CARE's more ad-hoc approach

It is also important to recognise that NGO approaches have changed considerably over the last decade to more sustainable models (microfinance is itself an example) and that the appetite for them also develops as part of this wider system of change. FSD's engagement with CARE's SG programme took this in a new direction because it sought to find delivery mechanisms that reduced the cost of delivery sufficiently that a scaled-up programme would create a self-sustaining methodology. This could operate at the level of providers (i.e. a network of sustainable fee-for-service providers), or at the level of users with a sufficient understanding of how groups operate accountably and transparently that spreads through peer-to-peer learning. Indeed this latter approach would necessarily suggest a non-linear impact pathway with the expectation that network effects will produce a 'tipping point' into an irreversible shift in understanding and the sustainability of user-owned groups. How to keep the process alive, adaptive and learning during scale-up is an important issue to keep in mind.

3.3.3 Was the context complex?

In the story of the history of the project, there were many examples of how the **history** and **context** were relevant and shaped the situation 'on the ground' – e.g. move from Rusinga, impact of previous collapse of SACCOS in Suba.

There were also **historical** factors with CARE itself that had shaped the expectations and views of 'best' practice. Savings groups had been implemented in Zimbabwe and there was a lot of pressure to adopt the design that had been perfected there. Customisation was frowned upon and the Kenya project was nearly stopped as the project team were initially regarded as not implementing the process correctly.

As we engaged with the focus groups of stakeholders, there were also many indications of **emerging impacts**, that had not been intended or planned for but seemed, in the eyes of several of the stakeholder groups, to be starting to emerge.[6] And these emerging impacts could, in the future have an impact on the core impact of creating financial inclusion for the poor.[7] This issue of emerging impacts is discussed in the next section.

3.3.4 Did the project design and project team respond to the complexity of the situation?

As with the credit reference project, there was a lot of evidence that the project design and the project team were responsive to the complexities of the context. For example, the project team **adapted** to the resistance of the Rusinga Island communities[8] and **experimented** with the use of community-based trainers. They also **customised** the approach from that developed in Zimbabwe to suit particular contexts.[9] There was some evidence that the project was **anticipating** the future in the sense of spotting emergent changes and developing a response[10].

3.3.5 Reviewing impact

The process of exploring the impact of the savings groups work proceeded in a similar way to that for the credit reference project through interviewing stakeholders. However, given its scale and nature there have already been a number of studies examining its impact and dynamics. This section therefore involves two parts. First, we review the secondary material for the insights this gives into the complex world in which SGs are operating and hence the way in which impact is related to these features of complexity. Second, we show how interviews with stakeholders in the field demonstrate how impact pathways which are both planned and emerging need to be explored in impact work. This discussion feeds into the methodological proposals made in the following section.

3.3.5.1 Secondary sources

There are now a number of studies of SG operations and performance in Kenya (both of CARE's work and of the programme by CRS at the Coast) and a more limited amount of data that examines the impact on individuals. However, a key limitation of reviewing material from secondary sources from a complexity perspective is the lack of a time dimension. The material involves cross-sectional studies as is usual for this type of work, rather than charting how change has happened for particular SGs in particular contexts with their particular histories and path-dependencies. This material therefore allows

6 For example, savings groups were starting to form clusters; there were examples of new enterprises e.g. bee keeping or pickling of vegetables and a growing interest in establishing cooperatives; more men were joining the groups and starting to shape their agendas.

7 In that, for example, if cooperatives were formed and failed, the savings groups would fail too.

8 And indeed had to adapt again when, later, these communities demanded that savings groups were set up when they found out how successful they were elsewhere.

9 And later customised it again when working in Marsabit.

10 E.g. considering how to support clusters, linking with work on marketing and providing financial management training, talking to banks about providing products that would support savings groups.

us to highlight the diversity of contexts in which SGs are working and the implications of these studies from a complexity perspective.

The reports demonstrate that there is "enormous institutional diversity" (DAI, 2010, p34) in SG operations and performance. The causes of this diversity are indicative but systematic influences are not clear. Local experience of group-based interventions - especially of financial groups in the form of merry-go-rounds, table-banking etc. – influences take-up and performance since many groups were formed out of prior existing groups. Indeed, self-help ideologies have a deep history in both colonial and independent government policy. This **path-dependency** operates collectively in relation to groups as it affects the character and knowledge of groups in the community as a whole. It also works more specifically at the level of the trajectories of individual groups. Reports show that groups also sought connections with other NGO programmes rather than striving for the autonomy that the SG model tends to assume is the objective (M. L. Odell & Rippey, 2011). In this they were themselves looking for synergies.

In these studies, there is evidence that **context** does appear to contribute to differences in performance. These particularly emerge across geographic locations with evidence suggesting that some of the strongest SGs are in Rachuonyo District. The characteristics that contribute to this are not well understood. They may come from a range of causes in the local context and the way the programme has interacted with it (e.g Through staffing, project relationships, and the way intervention modalities fit into this and so on). In these ways the performance and sustainability of groups themselves is a product of the local context and histories of past policy. It may also be contributed to by the way that the intervention relates to other projects and local events which may operate with positive synergies in some places and negatively in others (e.g. Where people have experience of failure).

The research shows that - for individuals - the way small scale savings and loans offer opportunities varies depending on local economic **contexts**. For example, in tea growing areas research suggests that the SG savings and loan repayments synergised somewhat negatively with tea production because it precipitated the sale of tea by women through the free market (*soko huru*) system. This brought lower returns than sales through the KTDA marketing system in which men were more likely to be registered (Coady Institute, 2012). In coffee growing areas access to SG finance supported engagement with more diversified sets of livelihood activities. SGs can therefore enable consumption smoothing both directly through access to finance for particular needs, or indirectly through the way they support cash flow in a range of livelihood activities. Women appreciate the increased autonomy that access to these funds offers while men appreciate women's contributions to household expenditure. SGs in some contexts fit into ideologies of "busy women" (Elliott, 2012) whose commitment to development is expressed through group membership. They also offer opportunities to demonstrate leadership and engage with local authority structures.

Synergies are also evident in the way that SGs fit with the assets, skills and social networks of the middle poor and particularly for women who are more able to use and benefit from them than men or poorer people (DAI, 2010). SGs can then also be **synergistic** in both positive and negative ways with other services. They provide competition with other MFIs and even banks on the one hand, while providing attractive ready-made groups with which some formal institutions now wish to link.

This brief overview of the evidence (a more detailed review is available in Annex 3) suggests that the variety of performance results from a wide range of underlying conditions. Indeed, at this analytical level, it is not at all clear which of these factors is most important in determining performance. In addition there are a range of factors which are not investigated in the studies, that could be hypothesised as affecting the dynamics and hence performance of groups; these include the influence of the skills and leadership abilities of trainers themselves; the role of group leadership; the role of group composition (e.g. Extent of relatedness within groups) and so on.

From the viewpoint of complexity this demonstrates well the way in which an intervention interacts with an open-system which has context, history and path dependency. It shows how the ability of the SG to function well is the product of these very micro-level factors in a way that potentially brooks clear generalisation.[11] In addition, this review demonstrates the methodological point that cross-sectional studies may well fail to reveal the synergies of context, history and path-dependency. It shows how these interact with the conditions on the ground that can result in a tipping point for a particular group into self-sustaining strong group performance or lead to failure. It is interesting to ask whether, if the studies had been designed more explicitly to investigate and uncover 'complexity', and had paid more attention to the pathways of development over time, such studies could have been more helpful in identifying specific trajectories of change experienced by different groups.

3.3.5.2 Tracing impact pathways

In the field we interviewed groups of stakeholders - SG members, village chiefs, government development officers and CARE project staff - in order to get their views of the developmental trajectory of the SGs. One difference from the work with CIS was that in these interviews there was much more discussion pointing to impact pathways. It was possible to gain a qualitative view of the importance of emerging pathways through keeping a tab on the number of times certain themes were raised. The impact pathways reported below resulted from an analysis of the interview material. They enable some reflection on the impact pathways planned for and anticipated by the project,

[11] This conclusion is also consistent with much other research on user-owned group dynamics, see e.g. (Bouman, 1995; Johnson & Sharma, 2007).

Table 2: Impact pathways of savings groups

Impact pathway	Evidence from interviews	Questions arising
A: Anticipated impact pathways		
1. Financial inclusion (Direct impact)	SGs have become established as a stable part of the social and financial landscape.	▪ Are savings groups growing and sustaining? Are they replicating spontaneously? ▪ Is the methodology working well (e.g. methods of paying and supporting CBTs)? ▪ Is it scaling up cost-effectively? ▪ Is the model going off track in any way or innovating positively?
2. Impact on incomes; wealth; vulnerability and consumption smoothing. (Final impact)	Respondents report a wide range of uses for the funds they get from SGs ranging from household expenses to school fees, small businesses and trading, furniture and livestock purchase.	▪ Are people better able to weather shocks in terms of food consumption, keeping children in school etc.? ▪ Are they more able to buy food, clothes and medicines for themselves? ▪ Is there greater accumulation of assets such as furniture, livestock, housing etc.?
3. Economic growth (implied in main impact pathway – but can be specifically monitored.)	Some evidence of increased economic activity; more people undertaking small scale businesses.	▪ Are people able to increase their incomes from current business and livelihood activities or are these saturated and producing low returns?
4. Economic development (could be implied by main impact pathway – can be specifically monitored.)	There was some evidence of new types of enterprises being taken up. Some of these were being undertaken in groups.	▪ Is there evidence of new enterprises e.g. making honey, pickling vegetables? ▪ Is there evidence of access to new markets? ▪ Are these new enterprises individual or collective?
B: Unanticipated / emerging impact pathways		
5. Social impact	Evidence that SGs offer social solidarity and support, especially to women.	▪ Are SGs developing solidarity and trust between people? ▪ How are they affecting relationships between men and women?
6. Collective action among savings groups.	There was concern in groups about how to handle increasing amounts of cash. Some groups were clustering together and their aspirations were expressed by members and local leaders to form SACCOs.	▪ What is happening as groups handle more money, cluster together, have aspirations to be even larger? ▪ How are they tackling governance and accountability? ▪ How do they deal with bad debts? ▪ Is there a change in leadership (from women to men) – what are the longer term implications of this?
7. Financial market development	Many respondents indicated a preference for SGs over MFIs and banks. Banks are interested in linking to groups and CARE was working with a bank to develop special accounts. (See Box 3 below).	▪ How are savings groups changing the financial landscape (e.g. impact on MFIs, behaviour of banks (e.g. set up local branches, offer interest-bearing current accounts)?
8. Financial capabilities	Respondents reported that SGs are supporting a culture of saving.	▪ How is the financial behaviour of savings group members changing e.g. are they maturing into use of other financial practices? ▪ Is financial literacy increasing?
9. SGs as potentially influential local institutions	There were reports of how SGs were on the radar of government, business and politicians.	▪ Are savings groups being seen as a political force, or as an economic force? ▪ Are they using their collective power? ▪ Are they seen as powerful by the private and public sector?
10: The impact of savings groups on NGO approaches.	There was indication that other NGOs are considering the use of SGs as a core component in interventions.	▪ Is the SG approach being used as core to other development related activities e.g. for disaster risk reduction programmes? Is it changing belief in creating non-dependent aid and development?

and those that are not anticipated but about which indications were emerging. This material then suggests questions for follow up impact work.

The above table highlights the many dimensions in which SGs are interacting with an open-system that open-ended conversations with members and stakeholders produce.

There are two main points here: first, from a complexity perspective, the way in which these impact pathways interact with each other may be critical for the core anticipated impact pathway in order to operate successfully. The SG approach has stripped down the model into one which seeks to deliver financial inclusion alone – it has removed any linked services and (usually)[12] assumes that groups will be able to continue to operate effectively on their own after a year of training and support. Nevertheless there was much discussion about how people could in fact develop their livelihoods. There is huge aspiration for groups to support economic advancement of the members individually and collectively.[13] At the individual level, whether or not the local economy is supporting livelihood opportunities dynamically (economic growth and development) and the need for skills and assets to take advantage of these is of course critical to the survival of the groups themselves. SGs are dependent on these dynamics even if core interventions do not wish to engage in promoting the development of new market opportunities and niches.[14]

There was also evidence both in the existing studies and the field research that groups undertook collective activity of benefit to them. Moreover, groups were starting to cluster together in places visited – allowing them to establish, for example, a social fund to help those in difficulties and allow the potential to invest in larger enterprises. While programme staff knew this and could see both the potential advantages and disadvantages, the programme had left the area and did not have a follow up strategy to address it. [15]

The point here is that if successful, such pathways have the potential to support greater impact and if unsuccessful, they have the potential to severely undermine the core intended impact. Indeed, the risks of this are probably greater than the chances of success as widespread experience demonstrates. The danger is that the failure of those specific groups leads to more widely negative reputational effects (as has clearly been the case in the past). Indeed SGs could be at a critical juncture after a few years in an area as the dynamic effects of success and failure are likely to feedback into overall higher or lower demand for the approach. This could tip the demand for the methodology into a virtuous or vicious circle.

The second point is methodological, that through this identification of existing and potentially new impact pathways, it is possible to formulate further questions and seek evidence to gauge their importance and their mutual impact. This can be undertaken in part through reviewing secondary reports and also through on-going monitoring by the project team. They were able to see to what extent such 'weak signals' of possible trends continue to emerge – or indeed die away. This approach therefore highlights emerging trends (where there are still only 'weak signals'). It also identifies potential synergistic interconnections which may either be a threat or support to the core goal to increase involvement in savings groups in a sustainable fashion.

The weak signals of trends obviously provoke questions which need to be explored further. For example:

- To what extent are groups clustering? Should there be some support given in formation of clusters and federations and some information about the pros and cons of SACCOs? Could this growth in scale be a threat to the core SG process?
- There seems little evidence through these interviews of financial maturation into the formal sector currently. CARE is starting to engage with the banks to see if they can provide SG-appropriate products. This area of engagement with formal financial processes, and perhaps the need for new ones (which if SGs recognised their power they could possibly influence even more), is once again both a threat and an opportunity for SGs.
- Is the increasing numbers and power of men in the groups a threat or opportunity? Should this trend be monitored?
- To what extent are SGs leading on to economic and market development and to new enterprises? And to what extent are these group based or individually based? Does this pose a threat or an opportunity for the future of SGs (in the sense that if such enterprises fail, people can lose their money and 'blame' the SG process)? Should there be more monitoring of this development[16]? Should there be more links to organisations that can provide support to market and enterprise development?

To explore these and related questions about what is emerging in the ways groups are performing requires ongoing ways of collecting information about emerging trends. The approach that we piloted including interviewing groups about the project and the wider context and of signs of change, resulted in a wealth of qualitative material that can be analysed to show trends and triangulate from different perspectives.

[12] Some models such as CRS's Private Service Provider model enable on-going support to be bought as a service.

[13] See for example Krijtenburg (2013) and the ideology of "uplifting" through groups in the Kamba context.

[14] In fact CARE staff on the ground were engaging with other donors and interventions to develop connections for the groups such as building market links to Nairobi, developing new products e.g. pickling vegetables.

[15] Although it was much in discussion within CARE, and had been the focus of discussions in a recent meeting of SG actors. CARE had also provided some training to cluster leaders.

[16] Which CARE staff saw as an advantage to them as it gives evidence for future funding explorations.

This material can be used to trigger discussions between project staff themselves at their monthly staff meetings. They can discuss anecdotal evidence of emerging change, decide to monitor or collect further qualitative evidence, discuss implications for programme change or addressing other development needs. Equally, it may need fresh eyes from time to time (six monthly or even yearly) to meet with stakeholder groups and take an external view as we have done here. Where there is evidence that emerging pathways are seen to be gaining importance, approaches to monitoring them in more quantitative ways can then be developed. This forms the basis of the proposed approach to working prospectively that is outlined in section 4.

Whilst our initial (two day) research process highlighted where emerging pathways may interfere and affect each other synergistically or antagonistically, it would require further work to trace whether these potential synergies did indeed arise. We provided a snap-shot of this process but the complexity-informed emphasis on tracing over time underlines the need for designing an on-going process.

It is also possible to look for these trends by reading existing reports and speaking to other researchers with these questions of emerging trends in

[17] Which we reviewed but in less detail.

Box 3: SGs and financial market development. Information on emerging pathway derived from group interviews

One point was made by every stakeholder group – MFIs have lost their grip where savings groups have been established.

Villagers, government officials and CARE were aware that savings groups were reaching their limits unless they could find some way to deal with increasing amounts of cash.

CARE is in discussions to see how banks can respond – including offering interest-bearing group bank accounts. CARE was also talking about using M-PESA to act as a conduit to savings accounts in banks. They felt the banks were taking notice of savings groups since if they continue to grow, and they could potentially take a significant amount of capital out of the banking system.

One of the chiefs in particular was very keen on the idea of village banks. There was a general question (from SG members) as to why you would want to put your money anywhere else given interest rates were so much higher.

The development officers were very engaged when talking about changes to the finance sector as a result of SGs. They felt that banks were going to have to take note, perhaps be forced to set up local branches. They seemed very conscious of the collective power of the SGs to change things.

Quotes:

We would like to establish a village bank where we can keep the money from the SGs and also offer interest to others. For example, perhaps we could offer a service to the village school teachers so they can put their money in our bank and gain a higher interest rate than in commercial banks. (chief)

We have to have a bank account in order to receive our salaries but we then take it out immediately and put it into the savings groups. (chief)

No one in the savings groups has microfinance (savings group members and chiefs)

SGs are really impacting microfinance (CARE)

Banks are going to have to set up local branches. Lots of money is tied up in these savings groups and they want to get their hands on it (development officer)

We have worked with Equity Bank to offer products more appropriate for savings groups e.g. interest-bearing current accounts, group accounts, (CARE)

Mature SGs need banking services for savings (CARE)

95% have mobile phones and use M-PESA (CARE)

We are looking at how to use the phone as a conduit from SGs to the banks (CARE)

Agency banking is growing (CARE)

We are not into microfinance now (SG- women)

There is very little microfinance (FBO)

There is no microfinance amongst group members. Some people have bank accounts, no groups have bank accounts, M-PESA is used for transactions. CBTs

KWFT is dying. They don't come as much now. (Chief)

mind. And there is also the possibility of comparison with other projects, in this case with the work of CRS[17], to see if such comparisons throw up further questions or opportunities for learning.

3.4 CONCLUSIONS

The evidence from the case studies appears to demonstrate well that the contexts in which these projects took place did indeed reflect the factors indicated by complexity theory (as set out in Box 1). Both encountered aspects of history and context which supported or detracted from their efforts and had to find ways to adapt their intervention to these circumstances. Both also recognised the episodic nature of change and the potential for the emergence of new and unexpected features.

Both demonstrate how factors worked synergistically and how the projects had to seize opportunities, adapt and finds ways to act in synergy with these circumstances. In the case of CIS, FSD's interests converged with those of existing stakeholders to take change forward while also strategically leading this momentum into a new phase of full-file information sharing. These efforts in turn interacted with the wider system in the way rising interest rates resulting from macro-economic trends precipitated the concern to control them. This resulted in the 'tipping point' of mandating full-file sharing.

In the case of SGs, the factors are sufficiently diverse that it is not clear from secondary data, whether or how different aspects (local economy and livelihood activities; prior group histories and experience; local social structure and so on), are associated with strong or weak group performance and sustainability. At the level of individuals the requirements of the services, as shown in these sources, also interact in different ways with individual's assets, skills and opportunities. In terms of method, working with secondary data alone, which has not been designed to capture the time-evolution of change, was insufficient to allow exploration of questions of 'why' and 'how'. It identified contextual differences but it was not easy to unpick why these differences occurred.

Overall, the findings of the case study research suggest that the complexity worldview and complexity thinking is indeed relevant to FSD's context and programming: the contexts accord with the notions of complexity as set out in Box 1.

Moreover, this evidence started to highlight particular features of project implementation processes that were borne of this complex environment and showed the extent to which project teams responded to this complexity (as set out in Box 6). Both projects demonstrated clarity of vision and strength of intention, in terms of what was to be achieved. This included a credit information sharing system or appropriate financial services for poor and rural communities delivered through low cost mechanisms. It was clear that these strategic visions were held well within projects that were also involved in significant detail. For example, SGs involved the detail of training, incentives and supervising different delivery channels; CIS involved the detail of implementing data formats and getting compliance. Project teams did, especially in the early stages, learn to adapt their programmes to the contextual and historical factors they encountered. The CIS project also demonstrated well how opportunities to influence change could be, and were being seized.

Finally, the work piloted an approach to reviewing impact analysis retrospectively, through seeking the 'story' of the project, and the issues relevant in the wider context. This method works well where there are small numbers of stakeholders who know the project well as in the case of CIS. In the case of savings groups it was harder to have a retrospective discussion with the group members and local project stakeholders whose perspective on the present and future was more compelling. This instead produced evidence of 'weak signals' of potentially emerging impact pathways.

The work also piloted a method of identifying emerging impact prospectively. Interviewing with a view to identifying emerging impact pathways would seem to offer an effective method to engage with the ways these pathways start to emerge and interact. This can be done by establishing processes to capture information from staff in the first place and potentially also establishing six-monthly or annual group interviews to take an external, more detailed review. This could lead on to new hypotheses to monitor - for example whether groups that establish group enterprises do better or worse than the average, or whether groups where increasing numbers of men take leadership change their focus and aspirations.

Chapter 4

IMPLICATIONS FOR PROGRAMMING, EVALUATION AND IMPACT ASSESSMENT

In this section, we consider the implications of adopting a complexity worldview – for programming, for evaluation and for impact assessment. The third section in particular builds on learning from the piloted methodologies in the case studies.

4.1 IMPLICATIONS FOR PROGRAMMING

In Box 4 we set out a complexity-informed view of the nature of change. In summary, one cannot assume that linear cause-and-effect chains exist and it is necessary to question the implicit assumption that contexts are stable and that outcomes will indeed unfold from the plans we design.

The question addressed in this section is what does this complexity-informed view of the nature of change imply for methods of programming and implementation?

At its extreme, complexity theory could be taken to imply that there is no point in planning or monitoring because the whole focus is about dealing with uncertainty and adapting to changing circumstances. We would not take this view, however. For large-scale interventions such as savings groups or hunger safety nets, being clear – as to goals and methods and monitoring – and persisting in their achievement, remain of central importance. What changes (with a complexity worldview) is how to plan for a greater degree of flexibility to respond to the unexpected, to seize opportunities and to adapt to changing circumstances. It is also necessary to be able to spot both anticipated and unanticipated emerging signs of change by embedding processes of monitoring into the project.

Box 4: A complexity-informed perspective on the nature of change

- Systemic and synergistic:
  - There are multiple and interacting causes.
  - There are multiple and interacting outcomes and impacts.
- Historical/path dependency:
  - There can be time delays between interventions and impacts (and things may get worse before they get better – the so-called 'J' curve[18]).
  - The impact of past interventions and events can affect the impact of current interventions. A conventional baseline may not capture such effects.
- Context
  - Change is heterogeneous, is affected by the local context.
  - The context itself is dynamic and changing during the project.
- Emergence
  - Interventions may have unintended consequences.
  - New features can emerge that were not expected, planned for or monitored.
- Episodic:
  - Change is episodic – periods of relative stasis followed by periods of radical change with emergent features that could not have been predicted.

A complexity-framed approach to programming is summarised in Box 5. The first step is to undertake a broad **context analysis**, in which time is spent both researching the wider context – the social, political, economic and environmental factors and the role of key actors within it – and **taking a historical perspective** on these factors, to understand what has shaped the current context and what is the strength of current institutions and patterns and features.

In parallel it is important to undertake an **organisational analysis**, to understand the strengths and weaknesses and attitudes and culture and institutions of the funding and implementing organisations. This will make explicit how embedded attitudes and institutions might affect decision-making, what will help and hinder working towards the goals and what work should be undertaken directly and what through alliances with other actors.

The next step is to **'weave' a set of goals or intentions,** working with a broad stakeholder group to agree what can be achieved - informed by the context analysis to take account of these myriad and systemic factors. 'Weaving' is taken to signify a process of reflexive and interconnected discussion from which new ideas can emerge. This is in contrast to a process of arguing from distinct perspectives, where outcomes are more influenced by which argument or which stakeholder has most power.[19]

Such a group can also be established as a steering group for the project, able to review progress on a regular basis, and debate where changes – to intentions or to implementation - should be made along the way. The idea that things will not go to plan, that the unexpected will emerge, that the context will change will require a different sort of budgeting, providing contingency to allow flexibility. It will also need a different sort of mindset.

Then the **initial steps** of the project are planned and **piloted** (or more than one approach may be piloted), to gain insight into what works best

[18] Woolcock, 2009.

[19] So, for example, if there are social arguments standing in opposition to economic arguments, the economic arguments often 'win' – a reflexive process is more likely to find new solutions which are 'win-win'. This is treating the process as complex – allowing multiple interconnections and providing the space for emerging and innovative intentions.

and under what circumstances. Based on the learning from the pilot, a wider implementation is then carried out. But the difference here is that there is a strong recognition that approaches may still need to be **customised** for different circumstances and need to **adapt to what emerges.** Ideally, detailed plans are still relatively short-term within the framework of longer-term plans which are more sketched out. There needs to be regular review and reflection to see what has worked well and less well, what requires changing and what has resulted or emerged that requires attention. This review process is best handled in a deliberate fashion so that any learning is explicit and decisions are well-justified. Otherwise, implementation can be too uncertain and provide an excuse for lack of persistence.

In general then, a complexity-informed approach to implementation finds a new middle ground between pre-planning, analysis and persistence on the one hand and agility, experimentation and adaptation on the other. This is in contrast to traditional approaches which tend to emphasise the former over the latter. A complexity-informed approach requires shorter cycles of planning and review to respond to the dynamic nature of change and the potential for changing contexts. It emphasises that working a wide stakeholder group, taking note of the views of those nearer to and further away from the project and embedded in the local context with those with a wider view – is more likely to recognise what is working well and what is not, what is changing both narrowly and in the wider context. It requires attention to the wider systemic contextual issues both before and during implementation; planning and implementation become more reflexive.

These ideas are not new and would be considered good practice in much project implementation – but a complexity worldview stays the hand of those who would cut contingency budgets, those who would feel that there is a 'best practice' approach that can be applied universally, those who would expect things, once planned, to go to plan and who would feel that implementation can be undertaken without too much attention to its interconnections with wider systemic factors.

It is interesting to note that the steps set out here were in general followed within the case studies. The credit-information-sharing project followed these steps instinctively (in the sense that the PM and key FSD staff tend to work in this step-wise way, sensitive to new challenges and to what unexpected outcomes and events must be faced). The early days of the savings groups 'story' illustrate the struggle between a PM who was sensitive to adapting to context, experimenting and adapting and an organisation who had more of a sense of a 'right way' and an optimised methodology they wished to be rolled out. The FSD approach has been to pilot options and closely evaluate their effectiveness – and the complexity view emphasises the need for continuing review and looking-out for the unexpected.

It is also important to point out that *when* to persist and when to adapt - or *when* to adopt tried and true models and when to customise - are matters of judgement. It is very important when making these calls of judgement, to work with a broad stakeholder group which includes representatives of those working in the field. It is also important to allow a level of autonomy and not over-manage at too high a level. These judgements often need to be made when the evidence is qualitative and emerging but is nevertheless tangible to those close to the situation, who have built relationships at the local level and are in tune with the developing issues.[20]

How can learning about adaptation to contexts best occur? Major differences to context are easily recognised – such as taking the SG methodology from Western Kenya to Marsabit. But how adaptation is needed within what appear to be more homogeneous environments is often less well understood. The diagnosis of uneven project performance is often treated as a learning feature in the *early* stages of project implementation. However, the search for understanding what works becomes less patient as the project timeframe evolves and there is pressure to achieve outcomes. Recognition that effective adaptation to context and local conditions is necessary is increasingly recognised by approaches which contract to pay by results as they allow greater flexibility of what and how implementation takes place. Of course this produces many hazards as to how incentives for results are constructed and the bias these create with respect to quality. This is an issue with the different SG delivery systems that the SG project is reviewing.

In summary, the complexity perspective would bring attention to the following issues:

- Expect uneven performance over time.
- Find ways to work with/address the unexpected and unintended factors that work towards/against your goal - the things that are beyond the risks/assumptions that can be defined a priori.
- Recognise and learn from variation across regions and between regions.
- Work with and monitor wider impacts where possible in order to understand feedback loops and how these also support / undermine the goal.

With projects that are more unique, more a 'one off' as with the credit reference project, then the balance between programming and seizing opportunities and adapting will differ from a larger-scale roll-out as with savings groups. The importance of maintaining and bringing into the foreground the long-term goals and using these to make choices about where to place effort, cannot be over-emphasised. Otherwise projects can wander in an unfocused way. Having said this, there will be aspects of such strategic projects which will still

[20] The literature on "brokers" and "translators" which explains how field workers operate to develop relationships and interpret project intentions and support into local contexts, along with a literature on street-level bureaucrats which explains how policy is mediated by front line implementers is relevant to this issue (Lewis & Mosse, 2006; Lipsky, 2010; Mosse, 2005).

follow conventional practices of programming; for example, implementing data collection standards, or designing and rolling out communication and training processes. But the decisions to undertake these may arise during the project rather than it being clear at the outset that they would be required.

Box 5: Programming to take account of complex contexts

1. Undertake a broad-ranging and historically-informed context analysis including review of current and previous interventions.
2. Undertake an 'organisational analysis' of funding and implementing organisations to identify locked-in institutional factors and lacunae.
3. Establish a broad stakeholder group and 'weave' long-term goals and intentions together, building on the context analysis.
4. Decide broadly on theory of change and theory of implementation in light of (2) and budget – but keep a contingency budget to allow experimentation and the need to adapt in later stages.
5. Pilot designed approach for the initial stages – or approaches – and refine methodology.
6. In light of 4, define programme, and theories of change and implementation and commence implementation; make explicit whether different contexts need differing approaches.
7. Allow for a degree of customisation to interpret intentions at local level but in a way that is documented and can be reviewed.
8. Develop detailed short-term plans but allow longer-term plans to remain looser.
9. Design in regular review processes (note, depending on the size of the project this process may need to be replicated at a more local level as well as at the overall programme level) – with a broad stakeholder group - which pays attention to:
   - Unintended outcomes.
   - Changes to context.
   - Signs of emerging impacts *including* those that were not expected.
   - Different experiences in different contexts.
10. Plan next stage after review in detail.
11. Modify programme and monitoring in light of 8 – but in an explicit way; document changes. Allow for responses to context but make explicit what has changed and the reasons for this.

4.2 IMPLICATIONS FOR EVALUATION

Complexity thinking requires a **path-dependent** approach that connects the past with the present and the present with the **emerging** future. So any project team or management team must be adept at looking forwards –be able to scan for changes and explore scenarios and 'foresight' (see footnote 2 above). They must be adept at looking backwards, at understanding what has led to the present, at identifying what deep-seated institutionalised factors are in play. And they must look in detail at the present, seeing the details and particularities of the **context**, both wide and narrow.

Complexity thinking also requires recognition of the **episodic** nature of change. This requires sensitivity to critical junctures which may become the moments for radical change to emerge, and a willingness to persist in 'building the ground' in times of stasis.

Finally, complexity thinking emphasises the dynamic and uncertain nature of 'the way things are' – the fact that there are inevitably unintended consequences and unexpected events and shifts.

How do we encapsulate these requirements into competencies? Box 6 provides a summary. There are competencies associated with forward-looking, with thinking of the longer-term which we have labelled as 'aspiring'; there are competencies associated with looking a few steps ahead, seeing where critical junctures might occur and working towards amassing influence and resources around such key opportunities. We label these 'anticipating'. There are competencies associated with dealing with dynamism and uncertainty – which we collect under the heading of 'adapt'. And finally there is the need to respond to the notion that every context is unique and requires a degree of 'customisation'.

The need for projects to be designed to scan for changes in the wider environment, to recognise unexpected outcomes related to the project and to adapt accordingly in line with the complexity of the situation, changes the focus for evaluation. Conventional evaluations of project implementation processes would not necessarily look for evidence of adaptability and experimentation (as set out in Box 2) but rather, would focus primarily on whether plans and actions defined at the outcome had been implemented according to the plan. Adopting a complexity-informed approach to programming therefore has implications for how the project is evaluated. It will be important to evaluate to what extent opportunities were seized, relationships built, experiments undertaken - all shaped by clear and strong longer-term intentions whilst not losing focus on outcomes.

Box 6: How should a project team behave to be successful in a complex world?

A meta-theory of implementation

- **Aspire**
  - Take a wide and systemic view of emerging trends – foresighting and scenario-planning.
  - Articulate long-term goals and intentions and use these to develop plans, prioritise actions, and determine which opportunities to seize.
  - Persist when there is little evidence of change.
- **Anticipate**
  - Think a few steps ahead during implementation. What might be the consequences of actions, or the consequences of events in the wider context?
  - Think through where critical junctures might occur.
  - Scan for changes in the context.
- **Adapt**
  - Adapt to unexpected events, unintended outcomes.
  - Seize opportunities, making choices in line with long-term goals.
  - Plan activities and projects but review progress regularly and take note of changing circumstances and be prepared to modify if necessary.
- **Customise**
  - Take account of contextual and historical factors in developing plans and programmes. Do not expect there to be a universally-appropriate best approach.
  - Experiment – pilot approaches, learn from what works.
  - Take note of the specificity of skills and relationships and factors in the wider environment.

Box 7: The FSD programme from the dynamic perspective of complexity

In its bid to enable change to happen in the financial market FSD's projects can be seen to embody a range of approaches which have the qualities discussed above. Examples are as follows:

Aspiring means that having a perspective on the future development of the market is vital to understanding how to strategize for change and interpret emerging developments. This involves having foresight which in turn requires constant and dedicated scanning of the wider environment, connections with key thought leaders and forward thinkers and the ability to construct likely future scenarios.

The scenario building exercises FSD has undertaken with the CBK and Ministry of Finance and summarised in the strategy itself, offers a vision of how change may or may not happen. The drivers for these changes that are identified offer indicators for monitoring to what extent different scenarios may be emerging in practice. FinAccess retail and business have also provided a knowledge base which feeds the ability to plan and develop strategies for improved outreach and effectiveness. However, change can take time and enabling it to happen is not always about being responsive and testing out approaches. It can also be about building the foundations of understanding and capacity in areas which may take time and effort. This means **persisting** while keeping the longer-term aims and goals in mind. Examples are where FSD has continued to work with the SACCO sector in implementing reforms and building capacity and with the FSAs through their transformation.

Anticipating requires attention to thinking several steps ahead, considering how certain actions or current trends may play out into the future. A number of projects in the portfolio are anticipating how the sector is developing and building the base for these to be effectively implemented as it moves forward. Examples include: work on financial education and capabilities; Government payments; the National payments switch; and health and life micro-insurance market analysis.

Adapting to the unexpected unfolding of events is something that FSD has repeatedly done and this consequently involves seizing opportunities when they present themselves. For example, the support to CBK in understanding the regulatory options for M-PESA as it started to demonstrate unexpected success (see Box 8) is such an example. FSD's projects often emerge out of opportunities although it is not easy to identify these from project documents. Further, FSD projects are frequently defined in sufficiently broad ways to allow for adaptation to occur as the project proceeds. The credit reference project was an example of seizing the opportunity of the CBK launching the guidelines for mandatory negative information sharing to leverage on this new development to lay the groundwork for positive information sharing.

Box 7 continued

Customising and experimenting with approaches and products is a feature of many FSD projects. Some involve getting something started in the local context (e.g. warehouse receipts and index based weather insurance); others involve working to get particular approaches off the ground (Centre for Branchless Banking; supply chain trade finance support; value chain finance centre; G2P payments development). When customising approaches to the context, it is important to do so in a rigorous manner so that learning can be captured and acted upon (and there are examples e.g. with savings groups, where FSD takes this approach).

4.3 IMPLICATIONS FOR IMPACT ASSESSMENT

In order to discuss the implications of complexity theory for impact assessment it is necessary to address the core of the current debate around impact assessment (IA) methodologies and this relates to the ways in which inferences of causality are to be established. Further background to these issues is provided in Annex 1. The following section then raises a number of issues related to the dynamics of change that this approach raises for impact assessment. The third section outlines an over-arching approach to impact assessment which spans 'working prospectively' and 'working retrospectively' that seeks to incorporate these insights. The final section briefly discusses how this approach relates to FSD's work.

4.3.1 Addressing causal inference

The key requirement of impact assessment is that it is necessary to have a robust approach to inferring causality from the intervention to the results. Weak approaches simply examine the co-presence of the cause and the effect, and many past impact assessments have done this. But a key result of the recent attention to randomised control trials (RCTs) and its reputation as a 'gold standard' methodology has been to raise the bar on the rigour required in assessing causality.

There are broadly two approaches to examining *active* causality.[21] The first is experimental and makes use of opportunities to manipulate an intervention such that people who randomly receive an intervention can be compared to those who do not (the counterfactual). Designed in this way, RCTs allow the direction of causality from the intervention to the impacts to be deduced.

Second, are approaches which examine the causal pathway in depth in order to construct or refine the theory of change.[22] These approaches are referred to as theory-based impact evaluation. Their standard of 'proof' is one which seeks to establish these connections "beyond reasonable doubt" (White & Phillips, 2012, p. 7) by starting with the programme's theory of change and then examining rigorously, using logic, whether it was borne out in practice. Methodologies which take this approach include contribution analysis and process tracing as these are generally applied (see Annex 1 for more discussion). Another called realist evaluation, seeks to identify an underlying theory by postulating mini-theories of different causal pathways and then examining which ones are operating in which ways in practice before drawing these together. The rigour of these approaches to causality is therefore derived from thorough and logical thinking rather than that of comparing experimental results against a counterfactual where there is no intervention.[23]

While RCTs have recently come to be seen as the most rigorous methodology in impact assessment, there are clear limitations to their use. They can only be used where there is a sufficiently large population of people or units of analysis that are affected by the intervention. So, for example, they can be used in principle to examine the impact of SGs[24] but not to assess the impact of CIS on interest rates. This has tended to draw an operational dividing line between the types of projects that can and cannot be assessed using RCTs. However, there are many operational issues that can also detract from their effective use (see also annex 2).

Other criticisms of RCTs arise from their inability to peer into the "black box" of what is actually going on inside an intervention. While it is necessarily the case that effects on which RCTs seek evidence are informed by a theory of change[26] the dynamics of how these are in fact operating are rarely revealed by RCTs themselves[27]. Results derived may therefore not occur as expected and be difficult to explain. Simply knowing whether average effects suggest

21 See Annex 1 by Befani in Stern et al, 2012.

22 See for example Vogel (2012) where a Theory of Change is seen to involve discussion of: context; the long term change intended; the process or sequence of change anticipated; the assumptions about how this might happen; and a diagram or narrative summary capturing this discussion. Her review that this needs to be kept

23 There is also a range of methods that use quantitative methods that are not experimental but which we do not discuss here. These are also frequently criticised as not attaining the rigour of RCTs, which include exploiting available datasets for natural experiments. They are also useful where data exists that can be exploited, but are not especially relevant to FSD's programme.

24 See for example (BARA & IPA, 2010).

25 Although they can be used to test whether the availability of credit reports affects people's borrowing behaviour.

26 For example, an RCT of SGs is likely to focus more on collecting evidence of whether consumption is smoothed or assets increased. It may collect data on whether participants have taken up local social or political positions but it may not go as far as collecting data on whether men have withdrawn or reduced their contributions to household budgets. Indeed, there are many potential effects and causal pathways and it is simply not possible to collect data that addresses all of them.

27 Though in theory they can be if the micro detail of causal pathways can be adequately captured in indicators such that those pathways can be appropriately tested.

something has worked or not may not adequately serve the purposes of interveners who wish to understand much more about the dynamics of how an intervention is working disaggregated by socio-economic categories and their interactions with contexts. It can therefore be argued that RCTs are, in any case, best complemented by investigation that seeks to examine whether the theories of change posited are in fact at work.[28]

Both of these approaches - experimental approaches and theory-based evaluations - assumes that there are sufficient regularities that causality is indeed identifiable. However, in the context of the approach to complexity outlined above, a more nuanced position on causality is to recognise that in reality causal pathways are rarely linear or independent. This means that what leads to what is always due to many factors at the micro level. If the context is stable and patterns emerge that are stable – then linear causality appears to be acting at the meso- or macro-level and is a sufficiently good approximation. But in general and to understand the dynamics of change, it is necessary to look at the detail - especially in the early stages of projects when evidence is weak and change is only beginning to emerge.

Complexity theory would argue that outcomes have multiple and interacting causes and indeed the ways these causes interact are likely to change over the course of the project. Seeking to identify what contributes to what is best achieved by following the actions and outcomes of the project as the project unfolds. If this cannot be done then the contribution of the project can best be investigated through retrospective pattern detection.

This is not at all to imply that working with a proposed theory of change when planning and implementing an intervention is not useful. But it does suggest that when approaching an impact assessment, starting from this theory of change may be too limiting to the approach to the enquiry and this would importantly be more open-ended. Approaching impact assessment is then about seeking to understand the *way* in which the intervention makes a *contribution*[29] to impact and to seek to identify the scales, scope and means through which this has happened. It suggests that in some contexts it may be more useful to do this in a more open-ended way to explore the *actual* pathways of change, than to theorise these in detail from the outset.

Even where causality is relatively well known and established through evidence from a range of contexts, (i.e. where RCTs might be an appropriate methodology), they may still need to be complemented by broader assessments. Impact assessment must always look for unexpected impacts. Some of these can occur both in the units of analysis specified in the theory of change and beyond it, as the effects ripple out to and interact with "wider" parts of the system. Such impacts may be particularly important where they operate in ways that importantly support or undermine the narrower theory of change. An example from the case of CIS (above), is the unanticipated use of credit reports for vetting public appointments. This may be important in gaining the acceptance of CIS by the public and overcome the backlash that was initially being experienced against the use of negative reporting. This wider pathway of impact invokes debates over integrity in public life and not solely individual relationships with lenders. Recognising and understanding this and its recursive effects on the intervention may be particularly important to the trajectory of the intervention in the Kenyan context.

As a result, this approach to complexity suggests that it is important to engage with the *detail* of the processes of change that are occurring, especially where the characteristics and degree of stability of the system[30] are not well understood. However, even where linear change has been evident in the past, this does not necessarily mean that it will persist into the future. This raises further issues for conceptualising and assessing the impact of interventions.

4.3.2 Further implications of complexity theory for impact assessment

There are a number of further implications of complexity thinking for conceptualising and assessing the impact of interventions.

Baselines

Impact assessment conventionally suggests that a **baseline** should be taken at the start of a project against which change can be measured. Complexity thinking suggests that it is necessary to attend to this in a richer way. While baselines are often established as a set of indicators of the context at a point in time, complexity suggests that it is important to understand it dynamically. In particular it is necessary to identify the way change has happened in the past and the contextual and historical factors that affect this. This would encompass particular social, economic and political factors that can affect outcomes and impact. This starts to lay a foundation for understanding the contribution of

[28] The debate over the impact of microcredit programmes and the use of RCTs is a striking case in point. Despite the limited evidence that high quality RCTs have shown for impact on incomes (see the systematic reviews on this subject), "randomistas" themselves have argued for the need to undertake further tests so essentially believing that the theory of change still stands and needs to be tested across a wider number of interventions. It might therefore be advisable to engage in more assessment that better seeks to understand why the causal pathways of the TOC do not appear to be working as proposed. From a complexity perspective, the point can be made that even though these appear to be operating in relatively stable contexts, the impacts they posit especially on incomes (or expenditure/consumption) are the product of open systems dynamics. These have such a wide range of influences that the causal pathways between microcredit as an input and increased incomes as an impact are the subject of many intervening dynamic processes. It is notable that many RCTs in this field are in fact conducted on much shorter causal pathways. This may be for e.g. between advertising and take-up of a product or between different modes of take up and delivery rather than variables whose causal links are necessarily more distant and therefore subject to more complexity dynamics (see for example (Karlan & Appel, 2012).

[29] Here we use contribution in the way used by John Mayne (Mayne, 2008, 2011).

[30] The relevant level of stability will vary. It could be a household, a village, a location / region or the whole country depending on the type of intervention and the focus of the effects sought. For example, the impact of access to finance for women can vary due to the state of intra-household relations. These might be stable and supportive and ensure co-operation in the repayment of credit or may be unstable producing conflict. The woman taking a loan may interact with this to exacerbate this situation and result in violence.

a project to processes of change that are already underway. This applies to baseline surveys of individuals, households or larger units of analysis such as sub-sectors, sectors, local or national economies. It suggests the importance of understanding past trajectories of change. This might, for example, lead to an individual or household survey that collects indicators of change experienced over a preceding period and then analyses them to understand the features of local economic change that are at work.

Emerging change

Following on from this, it is obvious that the context in which the project is being undertaken will also be changing during its implementation. When planning, project and programme managers usually consider what is happening and how these factors might play out in relation to the intervention. These factors are commonly captured as assumptions and risks in logical frameworks or similar planning approaches. Complexity underlines the importance not only of considering these at the outset, but of regularly tracking their evolution during the project to see whether and how they are building up or dissolving away, and what new risks or supporting (convergent) factors are emerging. While this is the terrain of good project management, it is not usually undertaken in ways that ensure these are documented and hence are available to project reviewers and impact assessors. This on-going perspective on the wider context and on emerging changes within the sphere of the project will inform the project process itself. It will also allow discussion about emergent impacts and requirements for monitoring

Impact in relation to the episodic nature of change

These features are important to recognise, because of the episodic nature of change. In order appropriately to assess the contribution of a project to change, the way in which change is occurring is important to understand. Hence some projects may be 'building the ground' while others are fortunate to be there when a 'tipping point' is reached. Without this perspective it is possible to over-emphasise the impact of particular events or projects, which seemed to 'tip the balance' and under-emphasise the impact of periods which changed attitudes and established the conditions for change, but where there was little evidence to show. The impact pathway may also be one in which the indicators of interest get worse before they get better and the dynamics of such situations are vital to understand as failure may be misdiagnosed. Projects intended to produce change may have made an important contribution though not resulting in the change expected. This may be seen for example, in attitudes which can be difficult to achieve, or that it was necessary to counter trends that were not initially understood.

This does not open the door to condoning failure on the basis that change is not yet evident. Project evaluators must apply critical judgement as to whether the project was effectively implemented against the features set out above. They must then clearly demonstrate with well sourced and triangulated evidence the way the project was addressing contexts and patterns that were slow to change, how it met with entrenched interests and sought to address them and so on. Impact assessors must similarly demand robust evidence of the features of the context and triangulate evidence of impacts across space, time and other interventions. Impact assessors must approach the task with a strong understanding of their methodology and set out evidence for their conclusions in rigorous ways.

Time

Finally, complexity thinking emphasises the dimension of time – the need to trace sequences of events and trace how impact emerges is part of the view of things constantly 'becoming' rather than 'being'[31]. This is to be contrasted with a view of the world as primarily stable - where how things change is less important than what they have changed *into*.

This does also suggest the need for assessing impact to be undertaken over a period of time which extends beyond the project's timeframe, as change may take time to emerge.

External validity

It is important to recall that the issue of external validity (whether the results found in one project will occur when the same project intervention is implemented somewhere else with similar features) is relevant to all impact assessments, whether RCTs or theory-based approaches. The perspective of complexity makes this even more significant. Indeed, it makes the issue of external validity a burning question from the programming perspective right through to assessing impact.

Complexity thinking brings into question the extent to which programme designs will ever work in the same way in different places even when the features of the different settings seem similar. At the same time, approaches such as process tracing offer a way to test core methodologies and understand the way they work in practice. So, for example, the core methodology of savings groups has been refined through paying attention to how they are effective though experimentation and through observing many implementations. Providing there is some flexibility to adapt to context, these core methodologies have stood the test of time. Process tracing speaks to the question as to how and why things work. In some cases, this time-sequencing will show up transient factors which were central to creating the path towards success but which can get forgotten and overlooked if attention is only given to outcomes. So following the history of the project can throw up surprising lessons and refine approaches about how customisation and adaptation are better managed[32] - issues which tend to be overlooked in the search for what works.

---

31 A terminology used by the complexity theorist Prigogine.

32 It is notable for example that in the vast array of literature on microfinance there is little, if any, that focuses on the process of how to customise approaches to different contextual and historical factors.

Box 8: FSD's role in M-PESA's success: Contributing to a "tipping point" but avoiding the over-attribution problem

The FSD Impact Assessment report of 2010 identifies the way in which FSD's responsiveness to the regulatory issues arising out of M-PESA's development helped it to get off the ground. FSD was able to source the input of an international consultant to advise the Central Bank on the regulatory issues surrounding M-PESA and to indicate to them that they did not need to regulate it directly. Instead they could afford to take a light touch approach because it did not represent a systemic risk to the banking sector or financial sector more broadly. This could be identified as a critical juncture in enabling M-PESA to operate in a way that allowed it to grow.

In terms of the attribution of M-PESA's success, it is clear that this input came at the right time and was very important. It enabled key factors to come together though it was but one moment in the development story, and Safaricom had done a huge amount of work in developing the model, its agent systems, marketing and so on. In addition, factors such as population mobility, use of texting, lack of current cheap, convenient and reliable alternatives for sending money, also contributed (Heyer & Mas, 2011). Research has shown how the success of M-PESA is also due to the way inter-personal transfers within social networks are an important feature of the way Kenyan's manage their financial and other resources. M-PESA synergised well with this social context (Johnson, Brown, & Fouillet, 2012).

Hence, while Safaricom was building the ground, FSD, with its particular focus on regulatory issues, helped to tip the balance and enable a tipping point to happen. If the regulators had not been sufficiently confident to allow it to go ahead then it is likely that M-PESA's growth would have not been as dramatic. In this case it is very clear that it would be erroneous to attribute the whole of M-PESA's success to FSD's intervention. It does however, illustrate why it is difficult to assess that contribution. The regulatory decision may have been necessary but was by no means a sufficient condition.

Based on (Stone et al., 2010)

Nevertheless, even in apparently very similar situations, it is important to remember that local events and circumstances will *always* make a difference to the pathway and outcomes will never be identical in different circumstances. So it is important not to over-engineer approaches and to allow space for adaptation and customisation.

4.3.3 A proposed approach

Based on the above considerations two related approaches are proposed through which FSD can more systematically engage with an impact assessment agenda in the context of complexity. These draw on the core ideas presented above and develop the methodology of process tracing. The idea of process tracing is that an explicit chronology of events is identified - through historical accounts, interviews and any available quantitative data. This seeks to identify the causal links and pathways which lead actions through to impact. In its pure form there is still an assumption of linearity and independence of actions. In some cases the actual pathway is compared with the expected pathway embedded in the theory of change and traced from expected outcome back through implementation. However, we believe that this technique can be developed to include multiple causes and be used both prospectively and retrospectively to provide an effective means of impact assessment in projects in complex contexts as discussed below.

Working prospectively

In developing methods that look forwards, impact monitoring and process tracing can be embedded in the project. We illustrated in section 3.5 how insights into the impact trajectories of savings groups could be obtained through interviewing stakeholder groups – e.g. chiefs, district officers, trainers, savings group staff and so on – and analysing and triangulating key points that were made during the discussions. This method was used to indicate a number of emerging impact pathways which were not necessarily those identified at the beginning of the project as core but which nevertheless could grow in importance and potentially influence the evolution of key impact goals as well as provide further contribution to economic and social development.

The most effective way to try and gain evidence of contribution is for the project team to look out for signs of emerging change or changing conditions and to start to monitor such changes as implementation proceeds. This can initially be done qualitatively, that is, by regularly noting these signs and changing conditions in project meetings and documenting them through management systems. As these are noted some will appear to fade away while others may become particularly important. Decisions can then be made as to which to trace. These can be monitored by collecting more data about them – this may be both qualitative (types, nature, contexts in which they occur, relative importance) and then collected quantitatively as indicators are identified which can capture the scale of these changes.

So, for example, if in the savings group project, project team members start to notice an increase in groups themselves taking up particular types of group level enterprise, they may discuss this in project meetings and should the evidence for their emergence persist, they may decide to monitor their type and number. They may also take note as to whether or not groups undertaking group level enterprises fare better in relation to the core impact of sustaining

the group to ensure greater financial resilience. In other words, how does this wider impact of enterprise development support or undermine the core impact pathway?

To take an example from the credit reference work, the project team had various information and insights on the way credit reports were being used. These ranged from being used to ensure repayment to their use in a limited way for loan assessment to how they had become part of the selection process for public officials. It is possible therefore to more consistently note these in monitoring reports. The project was already recognising their limited use in loan assessment and developing another project which would address this as it clearly demonstrated a lack of capacity in the sector to undertake risk-based pricing in the future. The way they were being used for screening high profile appointments (in banks themselves, in public sector offices, in parastatals and use in other private company recruitment processes etc.) could be monitored by noting them in newspaper reports or anecdotes. Tracing how this evolves over time would provide good evidence of an impact that could be significant if not intended.

Thus such tracing of wider impacts becomes embedded in the project process thus providing a much richer source of information by the end of the project as to how change is happening and thinking evolving. Also, being alert to new factors and unexpected impacts may inform the project process and allow learning and response which will, of course, also increase the likelihood of positive impact.

It is also important to point out that project teams need to become accustomed to noting what is new and may be significant and that such a process can be aided through providing an external perspective from time to time. For example, as some of the savings groups started to build clusters it was noticeable that leadership of the groups and clusters was increasingly falling to men. This could be a very significant change – if men start to dominate the management of the groups how will this affect the success of the groups? Such a change may be easier to spot from the outside as such changes often tend to 'creep up' on project teams. So a short review process involving external actors which meets groups, works with other stakeholders and seeks signs of emerging change, may be an useful addition to an ongoing process embedded in the project.

Working retrospectively

Whilst changing project and management processes in the ways suggested above are important in order to 'embrace complexity' as the project proceeds, the question still remains as to how to assess impact retrospectively. As discussed above, we suggest the development of process tracing as a particular method which can assist with this in retrospective evaluations and impact assessments (Hughes and Hutchings 2012). The argument is that if you want to try and establish contribution, you have to follow the project process from its inception (taking into account the 'complexity baseline' as discussed above), through its implementation and then into its impact. This both informs why and how things happened as well as *what* happened and gives the best possibility of assessing contribution as you follow the project actions and aim to trace what they contributed.

The interview process we followed, tracing the history of the project and exploring the wider context and historical antecedents as described in section 3, is one way to undertake this. It is obviously important to work with a range of stakeholders who can tell the 'story' of the project, give their views as to what happened and why and discuss what the project contributed. These accounts can then be triangulated and other information – in project literature or in wider reports – can also be used in this triangulation process. It may also be possible to identify other data which helps to contextualise the efforts of the project and therefore support the understanding of its contribution, such as newspaper reports on interest rate changes in the case of the CIS.[33] And the focus of the questions is not only to ask respondents about the project, but also to ask them what else was going on that might contribute to or work against success – to ask them about the wider factors and about other actors.

For savings groups, which is a 'large n' project, we dealt with the retrospective view through interviewing project managers and other key staff who had been involved with the project over several years, knew its history and were comfortable with this sort of reflective process. This approach would need modification to work with individual members, with savings groups directly, with SMEs and so on. This could involve using similar narrative methods with individuals which seek to explore changes that have happened in their lives and teases out the various ways in which this has happened and the contribution of the intervention to this. Given that the past is often difficult to reconstruct in this way over longer periods of time, impact monitoring visits with a cohort of individuals (a qualitative panel) which occur for example annually, could offer deeper insights into the trajectories of change occurring in their lives and allow the contribution of the intervention to these changes to be assessed. Such a methodology would warrant further investigation (see Annex 3 for more discussion).

[33] For example, the CIS project is intended to contribute to lowering interest rates. Establishing causality here is extremely difficult when trying to directly relate the level of interest rates in the economy to the implementation of CIS. The CIS has - at best - the possibility of producing overall reductions relative to what they would have been, but the actual level of interest rates is influenced by much more volatile and significant macro-economic factors. Effective assessment of contribution needs also to take a view on the relative role of these wider factors while commissioning a full economic analysis of interest rates in Kenya is not entirely appropriate to the task in hand. Monitoring movements through for example press and other economic reporting and their discussion of related factors at work provides a richer context against which the outcomes of the project can be assessed.

4.3.4 Deciding what is appropriate in FSD's projects

The proposed approach outlined in the last section offers an over-arching framework for FSD to start to extend its internal systems for monitoring impact. It can also effectively feed into evaluation and impact assessment to be undertaken both internally and by independent outsiders. The approach to working prospectively is applicable to all of FSD's projects. The retrospective approach is also applicable to most of its projects providing there are interviewees who have sufficient involvement with the project over time that they can take a long view. FSD may nevertheless wish to engage with RCT[34] type studies where these are particularly appropriate. This might be when (analytically) the conditions are judged as sufficiently stable and where the RCTs are therefore likely to have the real potential to yield important contributions to knowledge.

It is important to raise the issue of resources. It is necessary to invest adequate resources in impact assessment work of any kind whether this is RCTs or approaches that are methodologically based on theory and use qualitative research methods. While the costs of RCTs are now reasonably well understood, the kinds of qualitative approach which theory-based impact evaluation represents also require adequate resources if they are to be undertaken thoroughly. Rigorous impact assessment necessarily comes at a cost.

[34] Other quantitative methodologies have not been extensively discussed here but there is scope for these especially in FSD's research role as a knowledge generator for the sector.

Box 9: Complexity implications for impact

Systemic and synergistic: causality at the micro level is not linear. It is more appropriate to trace **pathways** through which interventions over time contribute to change except where very stable conditions allow inferences from considering stable patterns at the meso- or macro-level.

History: it is necessary to consider the base line not only at the start of a project, but also to understand the antecedents to this baseline so that the role of instituted factors is understood and the playing out of previous interventions is taken into account.

History: retrospectively, process tracing can be used to trace the pathways from intervention into contribution, triangulating across multiple perspectives.

Context: We cannot in general assume external validity – i.e. that if it works in one place it will work in another. Complexity offers a richer approach to assessing what is required in external validity.

Emergence: for the future, project and management processes should aim to spot emerging outcomes as well as unintended consequences and also monitor the broader context and track and respond to changes.

Episodic: too much may be attributed to an intervention at tipping points and not enough during periods of stasis. Therefore conclusions should not be drawn too quickly and impact assessment needs to be undertaken sometime after the end of a project.

Chapter 5

CONCLUSIONS AND NEXT STEPS FOR FSD

Complexity thinking provides a critique of and challenge to the level of certainty that is frequently assumed in conventional methods of programming and impact analysis. The over-arching message is that we can predict, plan, and attribute with less certainty than we would like. Complexity theory provides a worldview emphasising interconnectedness and the potential for the emergence of new factors. It can lead to a change in emphasis between planning and experimenting, between efficiency and responsiveness, between clarity of focus and breadth of perspective. It suggests that it is necessary to be somewhat tentative in concluding what works, and to be more willing to experiment. It emphasises that it is necessary to build in processes of learning and adaptation which respond to context and changing circumstances. It brings into question whether there are 'best' methods or unique frameworks. It recognises that taking note of history, understanding context and identifying changes through the building of a strong stakeholder group, and judging when to seize opportunities or adapt may be equal in importance to careful design and planning.

This research has shown that FSD's work, in seeking to catalyse change and innovate in a dynamic and fast changing market, certainly sits within a complex world. And this work has shown that, in many cases, project staff are cognizant of this and naturally adapt and adopt behaviours that are congruent with complex contexts.

The report has offered an over-arching perspective on how FSD can take forward the insights complexity offers for its learning about effective programming for change. More detailed proposals about how these can be implemented are given in Annex 3. Broadly these involve:

First, inculcating into on-going projects, processes which:

- Look forwards into the wider environment and consider potential shifts and changes that might materially affect the project.
- Use systems mapping[35] techniques to identify interconnections between interventions, and how these interact both broadly with the wider environment as well as more narrowly within the intended goals of the project.
- Review progress with a wide range of stakeholders on a regular basis to identify emerging impact pathways, identify new factors to monitor, and consider modifications to the project process.
- Adopt a prospective process tracing methodology, to keep track of emerging changes and shifts and identify how these were handled.

When the research was presented at a workshop with theme leaders (see Annex 3), they were enthusiastic about finding ways to incorporate these ideas into their management processes. This could also be taken further by piloting the approach in a systematic way in at least two quite different projects to examine how they can work. This would involve:

- Setting up an on-going process with partner projects to develop ways to capture relevant outcome and impact information at the ground level by working with field staff and management.
- Instituting regular review processes involving stakeholders and ways to collect and analyse this information (e.g. using a 'narrative fragment' approach - see Annex 1).
- Developing a methodology to work with a panel of respondents in large 'n' projects using a complexity informed qualitative approach.

Second, it is proposed that FSD develop and adopt a retrospective process-tracing approach to assessing impact and evaluating projects. This would build on the methods used in the research described in sections 3 and 4. The aim, through interviews with a wide range of stakeholders, through use of reports and other material, is to build up a picture of the history of the project and a triangulated (and reflexively-developed) perspective on the contribution of strands of the project to impacts, both intended and unexpected.

Finally, evaluating projects as they are implemented can be undertaken from the perspective of identifying 'complexity behaviours' as set out in Box 6.

[53] Systems mapping techniques (e.g. as described in soft systems methodologies (Checkland, 1999) vary in sophistication from post-it notes on boards to complex software packages. The methods allow conceptualisation and representation of interconnections between factors.

Annex 1

IMPACT ASSESSMENT AND COMPLEXITY THEORY

1.1 IMPACT ASSESSMENT: THE STATE OF THE ART

Over the last ten years, the elevation of randomised control trials (RCTs) as the gold standard for impact assessment in international development has provoked intense debate to the extent of being termed "paradigm wars" (Hughes & Hutchings, 2011, p. 1). The field of microfinance has been one in which such debates have been prominent. The ability of RCTs - or experimental methods - to address issues of attribution by ensuring internal validity for 'large n'[36] projects at the micro level, is achieved through the random allocation of the intervention between treatment and control groups. However, in practice, the gold standard is met by few impact studies in this field to date (Duvendack et al., 2011) with the difficulties of implementation raising a range of practical and methodological issues. Average effects in particular belie the heterogeneity of impact, reducing the potential for understanding for whom - and in what way - change actually occurs. Further concerns in use for summative assessment involve: the problem of low participation rates and whether they yield reasonable estimates; issues of variable treatment intensities leading to endogeneity and hence challenges in interpretation (such as the take-up of credit itself within a treated group); or the appropriateness of RCTs as an assessment method for recent, short-term interventions where results have yet to emerge (Bernard, Delarue, & Naudet, 2011). It is important to note that the issue of external validity - the applicability of the findings to any other context – remains, as for other impact assessment methodologies (Karlan, Goldberg, & Copestake, 2009).

Given that only some 5% of projects may even meet the criteria for RCT designs (Stern et al., 2012) this has elevated the discussion of appropriately rigorous methodologies and quality standards for impact assessment using alternative methodologies. These are particularly necessary where 'n' is small and qualitative methodologies are the only option. But the debate over RCTs has also exposed the limitations of non-randomised quantitative studies (Duvendack et al., 2011), and hence suggests the need for greater discussion of quality standards for these also. As Pande et al. show in a systematic review of impact studies related to financial inclusion, few actually interrogate or consider alternative causal mechanisms (Pande, Cole, Sivasankaran, Bastian, & Durlacher, 2012).

Going beyond experimental methods that enable attribution to the intervention, White and Phillips (2012) divide non-experimental methods into two groups: first, those that examine the causes of outcomes to establish "beyond reasonable doubt" *(ibid: 7)* how they have occurred; and second, those that are less concerned with establishing the causal process but for which participation of stakeholders and finding out their perceptions of the direction of change is vital to evaluating outcomes and are not regarded as sufficiently rigorous. In the first group there are at least four approaches: realist evaluation; general elimination methodology; process tracing and contribution analysis *(ibid)*. All of these involve identifying the explicit or implicit theory of change of the intervention; seeking to establish whether the outcomes have occurred as anticipated; and whether these have occurred based on the causes assumed or as a result of other non-intervention factors.

Realist evaluation focuses on understanding how the intervention mechanisms and their surrounding context produce outcomes in order to answer the question of what works, how, under which conditions and for whom *(ibid:8)*. It focuses on identifying, collecting and testing the evidence for Contribution Mechanism Outcome theories which are intended to build middle-range theories of how change happens. General Elimination Methodology looks at the range of possible causes and eliminates those for which there is no evidence of a causal link. Process tracing draws on evidence to connect an intervention to an outcome, and then identifies causal mechanisms to connect them and the evidence that would need to be observed to verify the actual process at work. Contribution analysis builds on this more process based approach to allow for reasonable inference -in contrast to attribution which focuses on necessary and sufficient conditions (Mayne, 2011). This argues that a reasonable contribution can be claimed if (i) there is a reasoned theory of change; (ii) that the implemented activities are those set out in the theory of change; (iii) the chain of expected results can be shown to have occurred; and (iv) other influencing factors have either been shown not to make a difference or their relative contribution has been recognised.

Evaluating against a theory of change, also called a programme theory, with impact pathways laid out has long been an approach to evaluation. Recently in international development there has been much attention given to the identification of theories of change as offering a means to improved programme planning as well as offering improved frameworks for evaluation. This can be seen as in part a reaction to the disparaging of the log-frame. The log-frame was originally intended as both a means of identifying a theory of change and the risks and assumptions that could enable or undermine it, resulting in an evaluation framework of indicators. This was derived through a process of stakeholder discussions. But its actual use became seen as an imposition by donors and poor practice resulted in "lack frames"; "logic-less" and "lock frames" (Gasper & Bell, 2000) leading to them being frequently regarded as unhelpful. The recent rise of theories of change seeks to re-establish many of the features that log-frames have lost. It seeks to develop programme theories derived from reflective development practice through stakeholder engagement. Undertaken in this way, the development of theories of change can expose assumptions and deeply-held beliefs about how change is expected to happen.[37] This allows sufficient flexibility for change to emerge and be adapted to during the project. Hence the approach

[36] That is where there are a large number of units - such as individuals, households or businesses - on which the intervention is expected to have an impact. 'Small n' refers to instances where there are few of these and hence RCTs and other quantitative methods cannot be used.

[37] There is reference to complexity thinking in these and reference to emergence, non-linearity, feedback loops etc. (see Vogel, 2012). However, how to work with these and what difference they make is not clearly identified.

can support innovation for programme improvement. This in turn demands performance management approaches that can allow flexibility in the context of uncertainty (Vogel, 2012).

In proposing an 'integrated' approach to 'small n' impact evaluation, White and Phillips suggest that: the questions of attribution must be identified; a theory of change should be set out; these together should be used to set out evaluation questions and assumptions to be tested and an evaluation plan devised for data collection and analysis accordingly; that alternative causal hypotheses should be identified and that the causal chain should be verified. However, this usage of theories of change tends to prioritise the establishment of causality over the process of stakeholder engagement i.e. it prioritises a search for objectivity over an acknowledgement of the existence of subjectivity. These approaches also take the view that people do what they intended to do; that there is stability in what was happening and that there is causality. They also act as if there is an external objective view which is somehow separate from the admittedly subjective, but also more granular view of those involved in the details – and time-sequences - of events and contexts. These priorities are established by the aid effectiveness debate more broadly in which objective knowledge is seen as being achievable and independence in doing this is required, so limiting the more participative and reflective dimensions. It highlights the tension between the relative power and objectives of stakeholders arguing for more summative externally-conducted approaches which seek to verify attribution and causality, over more formative approaches seeking to achieve adaptive programme management, and which are also more likely to be able to manage meaningful downward accountability to programme participants.

1.2 COMPLEXITY AND IMPACT EVALUATION

The implications of complexity thinking for impact assessment and evaluation have been considered by (Patton, 1994), Williams and Imam (*Systems Concepts in Evaluation; An Expert Anthology, 2006*), Stern (2012) and Woolcock (2009).

1.2.1 Classification of projects – simple, complicated, chaotic and complex

One trend in the literature important to programming, evaluation and impact assessment is to focus on the nature of interventions. Can the intervention be thought of as *simple, complicated* or *complex* (Glouberman and Zimmerman (2002), Rogers (2008), Stern (2012))? Woolcock (2012) and Snowden (2002) add a fourth classification, *chaotic*, to this list.

Woolcock (2012) describes typical 'simple' interventions as having predictable, simple causal pathways, giving the possibility of measurement precision. He suggests building roads or providing fishing nets are 'simple' projects of this nature. He contrasts these with the other end of the spectrum – chaotic – and suggests that projects focused on governance fall into this category. They have multiple pathways, many feedback loops and measurement precision is not straightforward.

Whilst this classification has gained traction, it still raises a number of difficulties. First, whilst with so-called simple interventions, such as building schools or latrines, outputs are easy to define and measure, this is not to say that their *impact* is simple. Are latrines actually put to use? How does road and pipeline infrastructure affect land ownership? How does such infrastructure affect the interests of the foreign private sector and how does that impact local livelihoods? Does it lead to economic divisiveness, an increase of shanty towns, single crop farming with degradation of habitat and reduced ecological resilience? Can simple interventions sometimes precipitate major and systemic change?

Equally, there is an interplay between the context and the intervention. Simple interventions in chaotic contexts will be much more difficult to evaluate than simple interventions in stable contexts.

So, whilst we would agree that interventions can be classified in this way, from the point of view of *impact* even the simplest actions can have complex effects, can interact in reflexive and complex ways with other interventions and with wider environmental factors. Conversely, complex interventions[38] can sometimes lead to effects which are simple to measure. The important question is to consider the interplay between interventions (complex or not) and contexts. In the work described in this report, building on (Boulton, 2012) and (Boulton, Allen, & Bowman, 2013) we focus more on analysing the complexity of the context (rather than the complexity of the project) and gauging its relative stability as a guide as to how to engage programmatically and how to approach impact assessment.

1.2.2 Literature on the implications of complexity

One theme in the literature, exemplified by Patton (2011), is to emphasise the *downside* of ignoring the inevitable non-linear interconnectivity – the complexity – of the context in which interventions are made. His thesis is that in general, environments or contexts are complex and dynamic, and thus interventions which take place in these environments initiate multiple pathways. What happens as a result of a planned intervention cannot, therefore, be determined with any certainty in advance – and assuming that plans will go to plan and planned outcomes can be measured and attributed to actions can reduce rather than increase effectiveness and/or lead project teams to 'hide' work that is not compliant (but which responds to the context) or to preference funding work (such as building schools) which is easy to measure.

Patton therefore argues in considering theories of intervention, it is best to explore possibilities, try things out, adapt in accordance with what emerges. In relation to impact, he emphasises that attribution is not straightforward, that

[38] Such as where a range of systemic factors are tackled together.

what emerges can either be triggered by the intervention, or be due to other factors, or be due to synergies between the intervention and other factors. His advice, in order both to determine the impact that is emerging and to be able to modify interventions to deal with unintended outcomes or unexpected events, is **to track emerging interconnections**[39]. He suggests that 'causality'[40] is to be understood as based on pattern detection, retrospectively constructed – not on trying to work out what causes what.

Patton is critical of counterfactuals, viewing them 'as meaningless because of complexity…. as there are far too many variables and possibilities emerging and interacting dynamically' (2010:24). He states that what needs to be measured can change during the evaluation as the process unfolds and new factors emerge and that it is important to track the forks in the road and implications of key decisions as innovation evolves. He emphasises that checking whether and how unanticipated and emergent features are dealt with is a fundamental function of evaluation. He advocates methods of *abduction* (inference to the best explanation) rather than deduction or induction, and contribution analysis rather than attribution analysis. Patton's central thrust is that impact evaluation is best thought of as developmental tool, a way to explore why and how change happens.

Woolcock (2009:2), equally, questions how often it is valid to make the implicit assumption that causes and effects can be linearly and reductively related. He argues: 'But it is only the most ad hoc theorising or wishful thinking (or the overriding imperatives of domestic political cycles and the structure of career paths at development organisations) that could possibly substantiate an assumption that all project impacts are linear and monotonic.' He gives as an example of non-linearity - the j-curve – where resistance to change can mean that 'things get worse before they get better', where, if you measured outcomes too early you would conclude that the intervention had had a deleterious effect.

Woolcock (2012) takes the example of a country establishing a new constitution. He argues that, on the one hand, it is possible and indeed straightforward to adopt a constitution from another country and regard the task as done. But does the constitution fit the cultural and historical and political context? Will it lead to good governance? Who owns it and believes in it? Will it change behaviour? Woolcock emphasises that development of governance and judicial and legal systems can take countries a long time, can only be achieved slowly and anticipatively and that our desire for clear and short-term outcomes and associated measurement can prejudice the undertaking of development. Instead, effective development practice and design needs to respond to the context, be experimental, seize opportunities, recognise emerging impacts, and adapt to changing circumstances. And, Woolcock emphasises, it needs to work with a wide range of stakeholders to try, in that way and ensure that the solution does reflect the relevant historical and cultural factors. Working participatively can also to ensure that the solution is 'owned' and hence interpretations of intentions during implementation are more coherent.

Complexity thinking supports the view that approaches to change and ways to implement will need to be sensitive to context and adapt to what emerges. This is not to disregard the importance of learning from other initiatives, and from being explicit as to how projects and programmes assume change is to happen. Rather it is to assess any theories of change and implementation as the project progresses and be prepared, explicitly and judiciously, to experiment, to seek contextual information and to make necessary adjustments. This 'middle ground' view is not well-developed in the literature.

1.2.3 Complexity-informed impact assessment methodologies

There is not an extensive literature in relation to actual methodologies for assessing impact with due regard to complexity. Some of the qualitative tools, reviewed by Hughes and Hutchings (2012) and White and Phillips (2012), start to move towards consideration of complexity, as discussed below.

Process tracing and contribution analysis

In particular, process tracing (Hughes and Hutchings 2012), seeks evidence for the chain of events that produce or contribute to a particular effect or outcome. This requires 'consider[ation of] alternative, competing explanations for the observed outcome in question until the explanations most supported by the data remain'. White and Phillips (2012) explain that the approach draws largely on qualitative data including historical reports, interviews and any available quantitative data to develop an explicit chronology of actual events, setting out the causal links between each stage. While adding time, there is still an implicit assumption of linear causality – that one thing can be deemed to cause another.[41]

While these approaches pay more attention to context than other methods do, they still lack a systematic approach to dealing with external validity[42] and still assume stability within the context and assume that a theory of change can be designed, articulated and can indeed be implemented as intended. There is the possibility, however, that some of these methods could be adapted to

[39] And this approach – tracking emerging connections – is what underpins our approach.

[40] Causality, in the physical sciences, has a very strict meaning and is used only to describe situations where a given input leads inexorably to a given output in a way that is predictable, independent of other factors and linear. The term 'causality' is used in this report in a weaker way to suggest strong evidence that (a) contributes to (b).

[41] Hutchings (2012a), though, has been experimenting with comparisons with 'actual' rather than 'espoused' theories of change. In other words the project team are asked what were their own intentions in undertaking certain aspects of work and then the actual work was compared with their 'local' theory of change.

[41] That is to say, complexity theory in emphasising uniqueness of context would argue that, because something works here is not sufficient evidence to suggest it would work 'there' – even if 'there' has seemingly similar qualities to 'here'.

be less predicated on a theory of change. They could represent a theory of change as something that can adapt to changing circumstances and that may be context specific. They could indeed, have the notions of adaptation and experimentation as the theory of change. The approach in this report, described in sections 3 and 4, seeks to trace the story of what did happen, but in an open-ended way, free of a theory of change.

The literature on narrative

Another body of work that is relevant to complexity thinking is the literature on the use of narrative as a qualitative research tool. Boje (Luhman & Boje, 2001) and Snowden (2002) and Byrne[43]), have both considered how narratives of events and projects can be used to capture complex dynamics. Snowden (see www.cognitive-edge.com) has developed software methodologies that allow analysis of 'narrative fragments'. The method is designed to work in the field, to allow respondents to tell stories and elements of stories and themselves work together to agree themes that represent and classify what these narrative fragments imply.

Stories people tell can indicate what people see as important, how they have made sense of events, what is the prevailing discourse and culture. Some qualitative researchers focus on collecting and analysing such collections of stories. Some researchers feel that such stories need to be 'complete', to have a sort of integrity, a beginning, middle and end. Other researchers collect fragments of narrative, capture fragments of conversation and would argue (for example, Georgakopoulou (2010)) that insisting on more complete stories prejudices the selection of accounts – preferences the stories of the more articulate, or preferences older, more well-established accounts which have been tidied up and embellished. Fragments of narrative, Georgakopoulou argues, are more immediate, more authentic and are less prone to bias in selection.

Clandinin and Connelly (2000:17) say:

'If we understand the world narratively, as we do, then it makes sense to study the world narratively. For us, life. . . is filled with narrative fragments, enacted in storied moments of time and space'.

The collection and analysis of narrative fragments formed the basis for our piloted process of identifying emerging impact pathways, as described in sections 3 and 4. The work described here only included one 'pass' at this process. To capture emerging trends would require more than one 'pass' spaced out over time, so it is possible to see whether emerging trends do indeed emerge, how they interact and what else happens that supports or hinders this.

1.2.4 Discussion – implications of complexity

Complexity has to be contrasted with a view of the world as machine-like. A machine-like view underpins most programming and evaluation processes. It suggests causes are additive not synergistic; that change is incremental and predictable; that outcomes can be mapped onto causes in a linear relationship; that change processes, once optimised can be standardised across a range of contexts. Complexity perspectives emphasise that this linear notion of causality and predictability is appropriate only in particularly stable contexts and in general to assume it is misleading and leads to unhelpful change processes and to erroneous methods of evaluation and impact assessment.

This affects the way implementation is approached. It legitimises a responsive, context-sensitive, flexible programming approach and strongly questions whether a theory of change can be fixed at the outset or is appropriate for all contexts.

Complexity thinking emphasises the path-dependent connections between the theory of implementation (how best to do the project) and the theory of change (the theory that defines how the project is designed to change the environment and with what intended outcome). The implementation phase shapes the impact phases and cannot entirely be separated from them so that discussion of impact inevitably has to return to the process of implementation, and seek to explore the pathway(s) followed.

When it comes to assessing impact, complexity thinking questions the efficacy and validity of defining a baseline. Even if a baseline can be defined, due to the time delay between inputs and outcomes, there is no way to be sure that subsequent changes in relation to the baseline were not caused by interventions which happened or were initiated before the baseline (rather than being caused by the intervention of interest). Moreover, impact pathways may emerge during the project – or indeed after the project is completed. And emergent impact pathways may interact with the intended impact pathway – impact pathways are not additive. So indicators determined at the beginning of a project may be insufficient to capture and track such emergent pathways. These pathways are not necessarily competing theories of change (although they may be) but components of a broader theory of change as they relate to the way in which this intervention sets off interactions within its environment – and how these develop are may be instrumental to the success or failure of the project and its core intended impact. Further, sometimes even the most important aspects of change may occur sometime after the project is completed, hence a need for impact assessment and monitoring to continue to some degree after the project is completed.

[42] Byrne, in a (Feb 2011) book proposal on complexity and social science says 'It will argue for a concentration on narratives since what is of interest to us is all of the past history, current condition and future potential of complex systems and any description of these has to be understood in narrative terms'.

This means that that evaluation is most effective when it is purposefully integrated into the project and programme from the outset. Indeed it is in the interests of projects to collect qualitative data on emergent features or signs of change as this can point to where more quantitative measures would be of value and where this can give evidence with which to seek further funding.

Complexity thinking thus questions the notion of attribution and instead points us towards analysing contribution. At a meta-level, complexity theory changes the theory of change. For example, if change is episodic, with periods where little tangible seems to change and moments where change can be very significant, this affects notions of contribution. There may be a tendency to over-emphasise the importance of the 'critical juncture' or 'tipping point' whilst under-emphasising the contribution to change of persisting and moving slowly forwards when there is little to show – but where this persistence got the situation to the point where tipping was possible.

Analysing contribution requires a degree of subjective consideration – no contribution is entirely separated from any other contribution. Since it is always a matter of judgement, it is important to consider from the outset who are the stakeholders and who engages in these judgements. Also, whilst in one sense *everything* and *anything* may be the trigger for change, not everything can be monitored. What is monitored is in part dependent on the vision and intentions of the programme and how these may change during the programme as evidence and new factors emerge. So, again, clarity is needed on who decides on these issues dynamically. This is not to suggest that complexity thinking questions the importance of evidence; rather it accepts that not all evidence is objective or quantitatively measurable. Moreover, it underlines that this relates to the power and politics of implementation rather than seeing the actors as in some way external to the theory of change.

Finally, it reminds us that even if we can demonstrate that a project or programme has had impact in one place, it is no guarantee it will have the same impact in another place (external validity). It also stresses the interaction between different scales of activity: micro, meso and macro. It is important not to reify levels and operate as if change can be understood as the interaction of patterns at one level (e.g. supply and demand curves). Qualitative change – that is, change which results in new features (development) rather than quantitative change (growth) which results in more of the same - cannot be understood by looking at the dynamics of relationships in one level alone. Indeed change to higher-level patterns is almost always impacted by the micro. So to understand the dynamics of change does require diving down into the detail at different analytical levels, in particular at critical junctures.

Annex 2

CASE STUDY - SAVINGS GROUPS

This annex contains a more detailed version of parts of the discussion in the text - in particular that using secondary data sources.

2.1 ACHIEVING OUTCOMES AND DIRECT IMPACT THROUGH GROUP LEVEL SUSTAINABILITY

It is clear from the secondary studies that there is "enormous institutional diversity" (DAI, 2010, p. 34) and experience of both group flourishing and failure. This next two sections use the available secondary evidence and material from our field interviews to illustrate how a complexity lens makes sense of this. The studies available deal with different projects (primarily COSAMO and COSALO1 with some evidence from SILC); they address different questions (group sustainability; individual impact; meso level impact) and use different sampling approaches and research methods (ethnographic; semi-structured qualitative interviews; focus group discussions; sample survey). This is not therefore either an exhaustive account of outcomes and impact but seeks to use this evidence to illustrate the underlying issues. Indeed it is not clear from the studies available what conclusions can in fact be drawn from this material – this is in itself of interest since it demonstrates a significant diversity of performance on which a complexity lens sheds light. The purpose in using this evidence here is **not** to make a judgement of whether this does or does not suggest that SGs are on balance a success or a failure. Such a judgement must be made against a set of criteria and a complexity lens is, in itself, agnostic about what these criteria are. Rather it points out that those criteria are determined by particular conditions (social, political, organisational, institutional etc.) and do not capture a complete picture, especially since they will be applied at a point in time. Indeed something that is regarded as a 'failure' at a particular point in time may then turn into a 'success' at another point in time or vice versa.[44] Rather if the issue is to understand the heterogeneity of outcomes and results across, space, time, social context and so on it is useful to analyse how this arises.

Past histories made it difficult to establish SGs in some parts of Western Kenya, such as Vihiga, as earlier pyramid schemes had undermined trust in savings institutions. Since some of these schemes had been sponsored by churches, working with FBOs to start SGs was a particular problem. Moreover, Odell and Rippey (2010) indicate that slightly over a half of the groups that they researched had a prior existence as a group in the form of a ROSCA, ASCA, CBO, welfare or self-help group or business. This highlights that the SG intervention may be one period in the life of the group rather than its start. They point out that it is necessary to ask whether the SG methodology might have undermined the sustainability of a pre-existing group. Indeed there is a much deeper path dependent relationship between SGs and a number of social and cultural institutions, and historical influences. The development of self-help extends back to the colonial period in which, for example, land terracing was undertaken with forcible sanctions and the movement of *Harambee* (self-help) since independence in the 1980s. The colonial government also promoted the formation of women's ROSCAs as a means of enabling them to have access to small sums of money, and these in turn also can be seen as based in the formation of labour groups among women. The relatively more limited operation of men in groups also has a socio-cultural history (Johnson, 2004).

SGs offer particular synergies with their context. For example, DAI (2010) points out that the share outs are usually timed to converge with the financially most demanding times of year (December – Christmas; February – school fees, planting; June – planting). The terminating ASCA approach of SGs solves a key problem that exists with ROSCAs and non-terminating ASCAs in that the share out enables **all** members to have a lump sum at a specific point in time. This is otherwise a reason why ROSCAs and non-terminating ASCAs can fail to serve people's needs as they cannot deal with covariant demand.

Nevertheless, the studies highlight the role of context in relation to performance reporting varied group sustainability across geographical locations. There are signals that group performance in Rachuonyo may be stronger than in other districts which raises further questions about context (Malkamaki, 2013; M. Odell & Rippey, 2010). A study of the meso-level impacts of savings groups (FSD/COADY 2012) does find that the agro-ecological context and the economic activities that arise as a result affect performance. In coffee and cashew-nut growing areas (eg. Malanga), SG loans were used to diversify economic sources. In tea growing areas in Nyamira, however, there was some evidence that SG loans were ultimately having a detrimental effect on household incomes and production systems, by encouraging short-term sales to *'soko huru'* traders rather than KTDA factories.

Odell and Rippey also document how about a 1/3 of groups have linked to NGO and related programmes of various types. This is unsurprising since many NGOs are looking for groups to work with. However, it underlines that the presence of opportunities to access NGO benefits may be both a motive for formation as well as a contributory factor to sustainability (see also Malkamaki, 2012). Additionally, operating as groups, SGs may wish to develop past activities or initiate new ones collectively to support themselves or others. There is evidence of both economic and social projects either in the form of income generation activities (e.g. renting out group assets such as chairs); mutual assistance in the form of social support or insurance; projects to assist children in the community with school fees being undertaken by 1/3 of these groups. This illustrates how synergy arises as well as the possibility that when groups have mobilised resources there are expectations from others for support.

[43] An example in the microfinance field in Kenya is either (i) Equity Building Society whose failure and virtual insolvency in itself set in place the conditions under which it was necessary for it to transform itself; or (ii) along similar lines the withdrawal of external funding from Partnership for Productivity in the 1980s which resulted in the managed ASCA model evolving in Central Province and developing a model with outreach to over 20,000 (Mule et al., 2002). Or the example of the Western financial crisis in 2008, that a booming banking sector– given time proved to be a source of failure for the sector as a whole – indeed part of a long history of cycles in this sector (find ref).

The SG methodology is designed to address the need to gain access to and manage financial resources in the context of uncertain livelihoods. It does this not only by creating a savings and loan fund within the group, but also through its ability to connect people to resources and processes beyond the group. It does this within a socio-cultural context which may be both a strength and a weakness: the long experience of group operations and self-help, and the political connections of groups have an impact on their success and failure. This illustrates how the sustainability of SGs is affected by the underlying features of a complex world. Many of these issues form the subject matter of social, political and economic analyses. The perspective added by complexity theory is the dimension of time; in particular how policies and interventions are affected by time and, as a result, produce synergies and interactions.

2.2 THE DYNAMICS OF IMPACT AT THE INDIVIDUAL LEVEL

This section reviews the evidence available on the impact of SGs at the individual level from a complexity perspective. The evidence from the DAI impact assessment study suggests that members were most concerned with the use of small scale loans available from SGs to enable them to manage their income generating livelihood activities both on and off farm (DAI, 2010). This contributes to the diversification of income generation streams which in turn lead to consumption smoothing.[45] Investment opportunities differ depending on the local economy (e.g. tea/coffee/cashew nut – see above) demonstrating the importance of synergy as well as context specificity. Loans which enable small enterprises can contribute to reduced dependence on poorly paid casual labour and allow women to undertake businesses that better fit with their domestic activities. While these contribute to income stability, given their small scale these activities do not necessarily produce new public market spaces for women (i.e. because of context-specific gender relations and gendered market development) and hence support (Meso-level study).

The SG pay-outs directly assist consumption smoothing and are also used for school fee payments and health emergencies. The availability of loans through this route can reduce the frequency of sale of small scale assets to cope with such eventualities. Availability of finance is also contributing to choices to upgrade schools for children as well as ability to reduce the time children might be "sent" (i.e. sent home) for school fees – that is, access to lump sums enable households to meet the demand for higher quality education. Share-outs are timed to meet particular financial needs - particularly school fees, planting and Christmas (DAI, 2010) .

The ability of women to contribute to household expenditure through small enterprises contributes to a sense of greater independence among women and it is perceived that "they don't waste anymore" (meso study p23) or are no longer seen as idle. The SGs fit with a local social context in which there is a "strong ideology" of groups. This also draws from past interventions that have worked with groups and which give rise to an ideology of 'busy women" (Elliott, 2012) whose commitment to development is expressed through group membership. In the case of SILC, groups have enabled women to create a new layer of leadership and local authority which is politically active - hence the opportunities provided by SG leadership are synergistic with the way local authority and leadership operates presenting an emergent impact. SG leaders also advise others on entrepreneurship and hard work. However, this discourse tends to screen out the role of men and stereotypes their irresponsibility which belies the reality of men's contributions to women's participation in SGs and the ways in which these new sources of capital synergise with household economies.

Membership of SGs is most easily taken up by the "middle poor" (DAI, 2010), Meso-level Impact study (forthcoming)) who have the social and economic, assets and skills to operate in social networks (i.e. the groups) contribute financially and manage group-based sources of capital (context and synergy with individual's socio-economic circumstances). Nevertheless, the GSL segment is more likely to be female with larger households and poorer than the average (DAI, 2010).

SGs create both competition and complementarity with other services (synergy). For some, membership creates credit worthiness in other institutions such as shops or schools. SGs are competitive with other MFIs and even banks, reducing the demand for their products and services (see also Annex 3 and Johnson et al. 2010). At the same time, they provide attractive new markets for formal financial institutions (Odell and Rippey, 2010). This is actively facilitated in some cases by SG programmes which work with formal providers to provide appropriate products to link with SGs and their members.

Finally, the creation of social funds in groups also enables management of emergencies. (Meso study, forthcoming). This provision of social support also creates social solidarity and mutual commitment (DAI, 2010) – which also fits with past experiences of welfare and self-help groups which are endemic to the local social context, past interventions and history.

This evidence illustrates the various ways in which the methodology of SGs and its results interact with existing features of the local social, economic and political context, including the way in which past interventions have worked with and consolidated group-based approaches to development. It illustrates how these are part of a system and how they therefore operate synergistically with these contexts. It is not possible to demonstrate from this evidence the episodic nature of change or how particular events synergise to enable change to happen (or not) since the studies available do not have a perspective on time in relation to processes of change.

[45] Of course, if the economic environment did not support small scale business this would not be so possible.

2.3 INTERCONNECTING AND EMERGING IMPACT PATHWAYS

This section demonstrates more broadly how the impact pathways of the SG project are interconnected and changing such that additional pathways also emerge. The core impact pathway hypothesised by the project as its theory of change is focused on the creation of groups to deliver appropriate financial services. It then looks at the impact of these services on individual's ability to improve incomes and reduce their vulnerability primarily through consumption smoothing. Studies so far have produced evidence of these effects, as indicated above. In this section, we use the evidence across the three levels (project, group, and household) and stages of the results chain (outputs, outcomes and impact) to identify other avenues through which impact is occurring. Evidence is also used to show how these other pathways interact with one another to synergistically disturb or reinforce this impact. The analysis raises issues about how we understand the heterogeneity of impact as well as raising issues for the sustainability of impact at all levels.

Financial inclusion and group replication: there is considerable evidence of the self-replication of the methodology (FSD Kenya, 2012; M. Odell & Rippey, 2010). This demand is most likely driven by the demand for small scale loans as evidence from other studies (including in relation to SGs) demonstrates (Johnson et al., 2012). So far the evidence on the sustainability of self-replicated groups is limited. Malkamaki (2010) finds no difference in the performance of project trained and self-replicating groups. Given the diversity of groups already in existence and motives for forming them, the factors leading to their potential sustainability are similarly complex (in the sense that term is used here). However, these factors clearly create the potential for both positive and negative feedback loops. Positive reinforcement of the value of the methodology where groups perform well emerge with negative reinforcement where they do not. Where groups perform poorly or fail there is the potential for SGs to gain a negative reputation and to create a risk to the system as a whole: poor performance leads to weak replications, and (like pyramid schemes) SGs becomes labelled with the risk of loss of savings.

Economic development: the evidence highlights the aspirations of individuals and groups with regard to the development of income generating opportunities. As access to small scale capital expands, so investment opportunities are subject to greater pressure. There is potential for market saturation since women (and other marginalised groups such as young people) tend to enter enterprise sectors where there are low barriers to entry. This constrains their ability to develop their income earning activities. This is especially the case when the economy as a whole is not dynamic. In this situation these enterprises are particularly saturated and business operators may be stuck in these and unable to move into new sectors and niches. The related issue is whether those with small amounts of capital are equipped with the relevant skills, networks, connections to markets and so on to take up such opportunities. The wider context of how economies operate and are changing is also therefore important for SG sustainability. Interviews in the field-work for this study suggested that some group members were indeed taking up new business opportunities, indicating the potential to support these pathways to impact both at individual and collective levels.

Collective action: some groups also sought to operate as collectives to promote their individual enterprises. This raises the question of how links to the support they need can be more systematically made and their capacity to manage collaborative activities developed.[46] There are also movements towards clustering of SGs, mechanisms for them to operate collectively or form SACCOs are related to the question of economic development as they highlight desires for resource mobilisation and development which is locally owned. This is something that holds the potential for strong impacts for some who are able to achieve it effectively but holds the risks of failure so often seen in such ventures in the past.

Financial sector development and linkages: the evidence also highlights the question of how SGs and their members interact with third party financial service providers. While some find their services more appealing than those of banks, the independence of SGs as financial service providers is by no means assured. Indeed they present a ready-made set of groups for the financial sector which increasingly wants to link to them. The question then arises as to what is needed to ensure this happens in the most effective way, as there is clearly a range of potential downside risks here that could undermine SGs.

Financial capabilities: there is evidence that SGs work well to support improvements in money management skills and the discipline of savings. This raises questions about whether and how there might be efforts to increase impact through the development of these capabilities. However, SGs can also be understood as supporting a culture of borrowing which offers challenges for capability development in terms of risks of multiple-membership and the need for juggling of debts which may not ultimately be welfare enhancing.

Social development: there is considerable evidence in existing studies of the importance of SGs in building social solidarity and social support mechanisms (DAI, 2010; M. Odell & Rippey, 2010), as well as support to women's increased autonomy and voice. They also ensure this support has a financial dimension through social funds, resulting in welfare enhancing impacts as well as material impacts on vulnerability. Social funds operating as insurance mechanisms, are dimensions of financial inclusion which are not given sufficient attention in impact assessments of SGs.

Group dynamics and local power relations: both the internal workings of groups and their connections to wider social structures and power relations

[46] There is a range of ways in which collective activities can take place which do not necessarily require the joint ownership of the enterprise itself or productive assets which, in the past, have been found to be very difficult for groups to effectively manage (e.g. posho mills). But collective activities such as joint buying/selling of inputs/outputs, organisation of access to markets and so on which support individual enterprises, are much more feasible.

are important to their survival and success. Malkamaki (2013) shows how these dynamics interact with local power relations, in which respect variations in delivery channels and social structures of groups produce heterogeneous outcomes. During our field work it was observed and reported by CARE staff that more men seemed to be taking on prominent roles in groups and clusters. It remains to be seen whether these trends are likely to support the development of the groups or create new challenges, since men have tended to find it difficult to participate in group-based finance in the past.

Political dynamics: the existence of groups as a forum for local politics also operates as both strength and weakness. There was evidence from our field work that groups are helpful to local administrations in their ability to mobilise people for *barazas* or other local events. (This a particular feature of the current election campaigns, also making it desirable to be a group member.) These engagements offer the scope to increase the voice of SG members in relation to local administrations and politicians, creating an alternative impact pathway to reduced poverty through the effects of political linkage on resource mobilisation and accountability for public service delivery. At the same time the incentive to join groups to gain access to external resources can undermine the core foundations for SG success, which rests on financial independence.

This discussion highlights the fact that once the "model" is devised it will interact with its environment in dynamic ways. As outreach grows new challenges are created, necessitating new responses to avoid negative feedback loops while continuing to strengthen positive interactions and impacts.

Annex 3

COMPLEXITY INFORMED IMPACT ASSESSMENT – PROPOSED APPROACHES

3.1 WORKING PROSPECTIVELY

3.1.1 Workshop feedback

The ideas of the report were presented to FSD's theme heads at an internal workshop on 1st February, 2013. This produced a lively discussion in which the ideas appeared to have substantial relevance for experiences of implementation and for the understanding of outcome and impact trends.

Theme leads were asked to discuss:

- Critical moments[47] at which the project significantly moved forward and in this context what were the ways in which (i) the project had built the ground for this shift and (ii) what else had happened beyond the project that enabled the shift to occur.
- Looking forwards to achieving the project goal – to identify what is starting to happen or might happen that will help or hinder its achievement?

These exercises immediately resonated with project managers who identified critical moments and what had been happening to shift the ground by the project itself and in the wider environment. For the second exercise, a brainstorming of factors that might help or hinder the project was mapped via post-it notes against the core causal chain of the project. This provoked an enthusiastic response as to the usefulness of this type of exercise for thinking through the evolving scenarios in which projects find themselves operating. While such factors are to an extent addressed as risks and linkages in project appraisal documents, they are not systematically revisited in project implementation in ways that are relevant for on-going management. In this respect, the consideration of the wider environment and factors justifying decisions to adjust course as projects evolve are not documented and captured in ways that evaluators or impact assessors can review. This vividly highlighted the disjuncture between models of planning and reporting which focus on a linear results chain, and the reality of implementation and impact in which projects are constantly responding to an evolving set of concerns.

Examples that came out of the discussion were:

- Financial Education and Capabilities – work on this area is proving difficult to conceptualise and move forward because thinking is still evolving and is therefore unstable. Decisions about what kind of strategy to implement in Kenya are therefore not easy to arrive at. However, there has been a strong basis of stakeholder collaboration and readiness to move forward with implementation. As the theme lead put it, his "nightmare" is that in three years' time the project will be evaluated against a set of global good practice and understanding which is different to the current thinking that is informing the project design. This suggests a stage at which conceptualising the field of work is itself chaotic and regularities in approaches and theories are themselves not yet clear. FSD is therefore engaged in trying to learn about what works and must recognise the risks involved and consequent implications for performance. It suggests the need to plan its own programmes in ways that can respond to its own emerging learning of what works and why and how it is working in the Kenyan context.
- Building capacity for the financial sector to serve SMEs (GrowthCap) – this project has been delayed by some two years due to difficulties in recruiting a suitable project manager. In this period the level of interest and understanding in the financial sector with respect to SME lending has shifted considerably with many banks now having strategies in place. This means that the portfolio of activities can now shift from a "push" to a "pull" approach. The landscape has also shifted as understanding from other FSD projects has filtered through. For example, the availability of credit reference reports has raised issues about whether the banks are using these appropriately in lending decisions tending to be used to produce a simple yes or no decision rather than used to develop risk-based pricing.
- The Financial Services Associations (FSA) project has persisted over many years to develop a model that could serve people in remote and poor rural areas with flexible and appropriate financial services. At a point where it is finally moving towards sustainability two wider factors can impact on it in completely opposite ways. The shift to County level government may offer new opportunities for the growth and development of FSAs in an environment which values local institutions. In the meantime on the other hand, the new development of M-Shwari (and its further evolution) may completely undermine their perceived value and the market for their services.

These examples demonstrate that FSD is operating in a sector that is undergoing rapid change in terms of concepts as well as in practice. It is doing so as the effects of policy and regulatory changes, together with competitive pressures to capture business ramify through the market. They also clearly demonstrate that the issues of attribution and contribution of impact to FSD are difficult to conceptualise and assess.

This session ended with some discussion of what might be feasible approaches to incorporating this understanding into project management systems. Two approaches were considered.

First, the newly introduced system of monthly project steering committee meetings was thought to be an appropriate forum within which this kind of reflection and scoping of wider factors affecting the project's progress could be embedded. To make this most useful it would also need to be captured in a record of the meeting. The maps that started to be developed in the workshop of wider factors and events contributing to, or constraining the achievement

[47] Also referred to as critical junctures.

of outcomes and impact could be a tool that was reviewed at the start of each meeting as a means of orienting the discussion of project issues. The 'map' could be reviewed and augmented for 15 minutes at the start of a meeting. This could help ensure that agenda items which deal with particular issues in project implementation are discussed within the frame of achieving the project goal.

Suggested questions to focus the discussion would be:

1. Wider factors that influence the achievement of outcomes and impact:
    - How have the factors that we have already identified as helping or hindering the project's outcomes and impact changed? What are the potential consequences for outcomes and impact?
    - What new factors are evident that may help or hinder the projects outcomes and impact? What evidence should we monitor to track the evolution of these factors over the coming months? (e.g. checking out through asking people informally before the next meeting; or more formally through a specific initiative to collect some data).
2. Evidence and signals of outcomes and impacts:
    - What signs are emerging of the outcomes and impacts that we planned for? Are these being effectively captured through our existing indicator reporting? How can we improve this?
    - What signs of outcomes or impact are emerging that we did not plan for – either positive or negative? What information do we need to capture these and to check out and monitor how important they are?
3. Connections
    - How are these emerging signs connected, how do these impact pathways - whether existing or emerging - support or undermine each other and the core impact pathway(s)?

Capturing these factors and noting them down in the minutes of the meeting means that they can be reviewed again for further developments at the start of the following meeting.

Second, a new project tracking format is being devised. This first involves theme leads identifying the theory of change of each project along with key research questions and possible qualitative and quantitative indicators. Then reviewing existing data sources and working with partners to track them on a regular basis. This will be stored in a programme database which will then be used to produce an annual impact report for each project and theme which can be reviewed by project managers and disseminated to stakeholders and funders. This could also be adapted to have sections on the factors that are contributing to the outcomes and impact of the project and those that are constraining it, answering the same questions as above. If adopted together with the project steering committee meeting it could act as a periodic and more formalised summary of the key issues that arose from the steering committee meetings with the indicators and evidence/indications which were collected also summarised.

Together these would offer a much richer body of evidence that would provide:

(i) A record of factors that led to the management decisions made.

(ii) A dynamic understanding of the evolving environment, events and factors that were either helping or hindering the project's progress towards its goal as well as frequent review of indicators that assess to what extent these goals are being achieved and whether other unplanned impacts are emerging.

These records would then provide a key resource to be given to evaluators and reviewers in undertaking Mid-Term Reviews, Project Completion Reports, end of project evaluations and impact assessments. Undertaking reviews of the project along the lines proposed in the methodology laid out above, offers evaluators a wider range of pathways within which to explore the role of the project. It offers the potential for a stronger understanding of the way in which building the ground has taken place and the factors that may have been detracting from this. For example, had GrowthCap in fact started two years ago and had to engage in more of a "push" approach to building financial capacity for the SME sector, the project might have encountered a range of negative forces which this type of impact monitoring exercise would have helped to identify. That is, when projects encounter frequent setbacks because they are operating in an area where the conditions are not conducive they may not succeed in achieving their goals but their overall contribution might be underestimated when viewed retrospectively. Similarly for an evaluator to understand the contribution of a particular project at a stage where the financial sector is more pre-disposed towards engaging with aims and objectives of the project, the evaluator should not over-attribute change to the project. Nevertheless, the ways in which the project effectively builds on existing dynamics can be assessed as strength.

3.1.2 Ways forward: piloting at project level

Given that there are the above potential ways to integrate a complexity approach into FSD's management systems, it would be appropriate to pilot the approach outlined above within one or two projects to see how it can work - especially given that projects are of quite different characters.

Prior to undertaking the steps proposed above it would be useful to review the project's causal pathway. This could proceed through the following steps:

Box A1: A proposed approach to reviewing, expanding and monitoring impact pathways

1. **Review and revise existing impact pathways and related indicators:**
   a. Review the core impact pathways laid out in project documents and revise if necessary to ensure these reflect the current best understanding of the planned results chain. Document any improved understanding of linkages and assumptions that has been gained from experience to date.
   b. Review the indicators that are attached to different points in the causal pathway for their relevance and usefulness. Propose and consider how to collect other indicators that are now understood to be relevant. Collect and review the actual indicators. If they cannot be collected without undue expense then consider what might be more appropriate, timely and useful to collect.
2. **Expand the view through undertaking the three step review process:**
   a. Wider factors that influence the achievement of outcomes and impact:
      i. How have the factors that we have already identified as helping or hindering the project's outcomes and impact changed? What are the potential consequences for outcomes and impact?
      ii. What new factors are evident that may help or hinder the projects outcomes and impact? What evidence should we monitor to track the evolution of these factors over the coming months? (e.g. checking out through asking people informally before the next meeting; or more formally through a specific initiative to collect some data).
   b. Evidence and signals of outcomes and impacts:
      i. What signs are emerging of the outcomes and impacts that we *planned* for? Are these being effectively captured through our existing indicator reporting? How can we improve this?
      ii. What signs of outcomes or impact are emerging that we *did not plan for* – either positive or negative? What information do we need to capture these and to check out and monitor how important they are?
   c. Connections
      i. How are these emerging signs connected, how do these impact pathways - whether existing or emerging - support or undermine each other and the core impact pathway(s)?
3. **Collect data:** decide what data / information needs to be collected on the wider related and emerging impact pathways:
   a. What information would better contextualise and explain the project's outcomes and the way wider factors are helping/ hindering these?

   For example, for the credit reference project relevant data may be:
      i. Sector PAR levels which have increased (Business Daily, 30th January 2013) because there are wider influences on this figure despite individual banks using credit reports to stop lending to problem borrowers and being able to use the threat of reporting to improve collections. Key indicators which capture these fluctuations are available from the CBK or the newspapers and the latter are also capturing on-going analysis of the relationships and trends at work.
      ii. Similarly, information on trends in lending is relevant. Such as new pressures to lend (e.g. due to an improved government fiscal position), new products and how these are affecting the environment for lending to SMEs and consumers. These are relevant to the way in which CIS impacts on the portfolio, methods of risk-based pricing and so on. These can be captured through tracking sector level developments through the newspapers and related analyses by the CBK.
      iii. Wider public awareness of credit information sharing can be tracked through newspaper articles on blacklisting and cases in court. Shifting coverage of positive information sharing among the public could therefore be seen through this discussion.
   b. Put in place responsibilities for collection of this information and reporting at the Steering Committee Meeting or through the new project reporting format.

4. **Review and decide:** once the data is becoming available, review the evidence and signals of outcomes and impacts against this data (step 2 of the steering committee meeting above):
    a. Discuss what this means for the project's impact pathway and goal. What does it mean or suggest should be done differently to achieve the goal? In what ways can a positive emerging impact be built on. [E.g. how could the use of credit reports to screen public officials be built on in the public awareness campaign?]
    b. Decide whether a decision can be made or whether more information is needed and how to get it and keep it under review.[48]
    c. Monitor to assess whether this impact or pathway is growing or subsiding, or fluctuating in a way that needs on-going monitoring. Consider how you can assess whether this represents an issue of growing concern or importance or has subsided and is no longer relevant and further tracking is not needed. How is it changing key patterns of behaviour? This may be done by checking out wider interpretations of importance beyond the project team, considering whether the impact or issue is starting to affect other areas of programming and hence is clearly becoming of broader importance (i.e. connections to other areas).
    d. Decide how the project needs to respond to this change. Document the decision and consider how it relates to the core causal pathway. Revise as per step 1.

Such an approach could be undertaken not just within FSD management structures but with project partners in project review meetings.

It is evident from discussions with FSD managers that it is this type of dynamic reviewing and constant re-consideration that is happening in FSD in informal ways. The purpose here then is to find ways to systematise it in management processes and capture it in documentation in order to make it more transparent to staff, management, evaluators and funders.

3.1.3 Implications for FSD programme level management

Theme heads were also concerned in the workshop that the interactions between projects and the conditions and issues they are encountering tend to be missed and not sufficiently discussed. The above proposed approach offers a means to identify these synergies and interactions. This could be done through a similar review process at the beginning of senior management meetings. It would involve a review of the project level factors that interact with other project pathways. A similar process could be undertaken for reviewing these driven by similar questions as at the project level.

1. Wider factors that influence the achievement of outcomes and impact:
    - What are the factors that are affecting the achievement of the outcomes and impact of the FSD programme as a whole? These are core to the issue laid out in the Strategy document of achieving different scenarios of cash lite or lock-in to a low- level development trajectory.
    - What are the factors that are affecting the achievement of outcomes and impact across more than one project? What are the ways in which outcomes are interacting between projects?

2. Evidence and signals of outcomes and impacts:
    - What signs are emerging of the outcomes and impacts that we planned for? Are these being effectively captured through our existing indicator reporting?
    - What signs of outcomes or impact are emerging that we did not plan for – either positive or negative? What information do we need to capture these to check out and monitor how important they are?
    - How are these interacting and connected across projects in the portfolio and for the programme as a whole?

These dynamics could also then be summarised into quarterly reporting at the programme level to FSD's donors in the quarterly and annual reports. This would give a more dynamic view of the way in which FSD is managing a portfolio of activities in a dynamic environment.

The workshop focussed on finding the relevance of complexity ideas for on-going management and for how this facilitates capturing the dynamics behind outcomes and impact. As discussed above, many FSD projects do focus on systemic change or work with a relatively small number of players, and hence many are not suited to more conventional approaches to impact assessment that deal with "large n" scenarios. The approach outlined above is a means of operationalizing management systems within a dynamic evaluation for

[48] One approach to ensuring that decisions are made and implications followed through into practice in a related but more focused area for managing responding to clients in MFIs is The Practice Guide on the Feedback Loop produced by the Imp-Act Consortium.

impact assessment. This does not mean that other methodologies such as RCTs, modelling or other quantitative methods are not also relevant to particular projects to answer particular types of questions. It does offer a basis for capturing the types of issues that FSD is regularly dealing with and trying to understand, in its bid to achieve its programme goals and hence impact on the financial sector.

The proposed approach for projects can be undertaken at the programme level and could form the basis of the proposed annual impact synthesis report as an annual review of programme performance which the Knowledge team will be producing in the coming months. The analysis produced from the above overview would offer a systematic review of the pathways to impact of the existing portfolio - or at least the major projects. It would provide a framework for the overall synthesis report into which other evaluations and studies would also feed. These other studies would also contribute understanding of how FSD's impact on the sector is evolving.

3.2 WORKING RETROSPECTIVELY

3.2.1 Evaluating retrospectively

Retrospective evaluation applies to projects where an evaluation or impact assessment is to be carried out – usually ex-post without necessarily having engaged in the above on-going approach. Where the approach to prospective evaluation has been implemented as proposed in the previous section then there will be a richer set of data with which to explore impact pathways and assess contribution through the approach laid out below.

An approach to evaluating retrospectively is given in Box A2.

Box A2: Evaluating retrospectively using stakeholder accounts of the project[49]

The retrospective methodology focuses initially on using and triangulating a number of accounts of the project or programme from different stakeholders and seeks to explore what happened, why and what was relevant in the wider context.

It builds on the notion of path dependence – how actions, outcomes and impacts build on each other over time. It explores the fine-grained detail of the implementation story. It does not assume a linear or incremental theory of change but tries to abstract how change actually happened.

The method has similarities to process tracing. While process tracing seeks to compare what happened to what the theory of change said would happen, this method does not impose a priori any theory of implementation or theory of change, but simply seeks to explore what happened.

Apart from the idea of tracing, what this method shares with process tracing is that progress is compared with the longer-term goals and aspirations of the project – what impact was intended, what the project was contributing to and working towards.

Data collection

To have rigour, it is important that the accounts come from differing types of stakeholders and players and ideally people from different levels and perspectives. It is clear that people to a degree tell the story from their own perspective, favouring their own agency and likely over-attributing to the project given that the origins of the evaluators are known. But triangulation provides perspective and offers new avenues of inquiry. Triangulation should therefore purposefully seek out views from those on the edges or with little direct interest in the project.

- This starts with questions pertaining to the 'backstory' – what led up to the project, what had happened prior to that, both in the context where the project was to be implemented, and in the relevant organisations (had similar projects been implemented elsewhere?). Discussions as to what was happening in the wider context and what other players were doing are also explored.
- The history of the project is then followed either to its end or up to the present. There is a chance to check details between interviewees, to explore differences of perspective.
- These verbal accounts are also triangulated with project documents and any other relevant documents.
- Final questions focus on impact – what had been intended, what was achieved (both in terms of outcomes and impact), what else was achieved that had not initially been intended.
- And contribution – how did the project contribute, how did particular features of the project contribute, what else contributed.

[49] This approach has similarities with the approach taken by OPM in the impact assessment of FSD's programme in 2009. However, that approach particularly involved discussion of a counter-factual scenario with project beneficiaries that is not the focus here. The approach then also involved getting views of a wide range of stakeholders. This approach emphasises the need to do this in ways that much more explicitly explore the patterns of change and factors supporting them and how the project does or does not interact with these.

Box A2 continued

Analysis

From these accounts, a narrative can be constructed which seeks to capture critical junctures in the project and also tracks the influences and dynamics of the periods where less seems to be achieved, but where the project is attempting to persevere in what we have called 'building the ground'.

From this narrative, there may be particular issues to explore in more detail, or particular hypotheses to investigate as to what was happening and why. These can be explored through further interviews, through returning to some of those already interviewed and also through exploring relevant written records. This 'second pass' is an important feature of the methodology and often results in new information and new perspectives.

It was interesting to note that respondents feel very comfortable in framing change as a mixture of phases: 'building the ground', moving forwards (or not) at 'critical junctures'. Use of 'critical junctures' has similarities to the method of capturing 'most significant change' – but it pays more attention to the way several factors come together – and it does not require these critical moments to be positive – i.e. they could be moments of collapse, moments where severe set backs were experienced. These can be explored in some detail where the dynamic approach to evaluation has been implemented, as a richer data set of conditions and emerging features surrounding the project's progress will be available.

The method is evaluative in that it traces the narrative of the project. But it also leads into explorations of impact - and can help to identify what quantitative measures of impact would support the assessment – e.g. if the aims of the project are in part to impact interest rates, it would be useful to obtain data on interest rate movements during the course of the project.

3.2.2 Complexity-informed qualitative methods for large 'n' projects

This evaluation approach suggests similar methods for considering how to assess impact where the unit of analysis is, for example, an SG, an individual or a SME.

Exploring impact at these levels would involve examining the trajectory of events, interactions and outcomes for the unit of analysis in question (individuals or organisations – e.g. groups, SMEs), including emerging, unexpected or unintended effects. This would be undertaken using primarily qualitative methods. An approach may be developed using the notion of critical moments or junctures to probe particular times which stand out in the minds of interviewees as of particular significance – where there were convergences and synergistic events and the potential for step change. The interviews can then reveal whether such moments were seized and a contribution made or whether the opportunities were either lost or where the hoped-for changes did not materialise. While this approach starts from these critical moments (and experience with this method has showed that interviewees in general have a very good recall of such moments and have no difficulty relating to the concept), the method also then allows further explorations to trace the slowly-shifting trends, interweaving of influences and intentions that might have built up to these points.

Interviews would be undertaken broadly along the lines discussed in the box above in which interviewees are asked to discuss how the intervention has 'built the ground' or tried to create or respond to 'critical junctures'. Specific exploration can also be carried out as to what else was contributing to them.

Purposive sampling can be undertaken on the basis of the existing understanding of the diversity of impacts occurring, since the purpose is to understand the diverse ways in which change occurs, for whom, how and why. This evidence can be used to assess the strength of a change in patterns –if the apparently weakest cases present clear evidence for a change in patterns this would more strongly support an overall assessment that patterns have changed.

In the situation of 'large n', for example, to assess what the impact of SGs has been on individual livelihoods, a proposed approach is given in the box below.

Box A3 : A complexity-informed qualitative method for investigating impact in large 'n' projects

- Use purposive sampling to explore the heterogeneity of impact - this might be e.g. in areas where the programme is known to have been more and less successful through its output or outcome monitoring data; within different levels of use of the programme; across gender, poverty levels etc.
- Undertake qualitative interviews also using purposive sampling (e.g. of SGs and their members who are known to be performing more/less well) which focus on exploring chronologies and critical junctures. Again using critical junctures as moments through which to explore the broader range of influences and events that were occurring and that contributed to this. This approach could take a two-way approach to cross checking the importance of the intervention:
  - First, ask in open ended ways about what has been happening in the life and livelihood of the individual over the relevant recent time period (e.g. two years). Explore these for critical moments and the trends and events that built up to and surrounded them.
  - Second, undertake a chronology of the individual's engagement with the intervention in question. Ask about critical moments in relation to this intervention and explore the issues around them.
  - The relationship between the two sets of critical junctures and their related conditions are then used to contextualise the importance of the intervention to the life and livelihood of the participant and analyse its contribution. [NB. This method has not yet been tested].
- Data collection at the context level such as the village through e.g. FGDs, and key-informant interviews would also explore the timeline around the intervention, what else was happening and what critical moments have occurred and what other conditions related to these at this broader level over the relevant period.
- Analyse this data for:
  - The outcomes related to the planned pathway and how these differ among participants using analysis of how their participation has been 'building the ground' for them as individuals or contributing to 'critical junctures' in their lives.
  - Identify the types of critical junctures that the respondents experience and also analyse whether and how these relate to unintended or unexpected outcomes of the project and their connection to other dynamics of change.
  - Analyse the extent of the relationship / interactions between the way 'critical junctures' and 'building the ground' have occurred in their lives, how these relate to the intervention and how they relate to wider processes of change.
  - Analyse the types / nature / conditions of 'critical junctures' and 'building the ground' in the relationship they show between the intervention and outcomes/impact experienced.
  - Once types of juncture / building the ground have been identified and the conditions in which they operate have been identified, analyse contributory causes using techniques such as qualitative comparative analysis to identify patterns of these.
- Draw conclusions about:
  - The conditions in which the programme operates.
  - The nature of change evident, for whom and under what conditions, and their relative importance in the lives of participants.
  - The nature of interaction and synergy between impact pathways both evident and emerging and the implications of this for the sustainability of impact into the future.

The findings as to the nature of junctures and building the ground and their conditions could then be fed into wider surveys or on-going impact monitoring. These would require the identification of indicators that appear to be related to these occurrences and indicators of conditions in which they appear to take place. This then has the potential to pick up particular types of indicator e.g. of behaviour change that relate to the use of the intervention and lead to the quantitative analysis of patterns of change.

REFERENCES

Allen, P., & Boulton, J. (2011). *Complexity and limits to knowledge: the importance of uncertainty.* In P. Allen, S. Maquire & B. McKelvey (Eds.), The Sage Handbook of Complexity and Management: Sage.

Allen, P. M. (1997). *Cities and Regions as Self-organizing Systems: Models of Complexity:* Gordon and Breach Science Publishers.

BARA, & IPA. (2010). *Baseline Study of Saving for Change in Mali: Results from the Segou Expansion Zone and Existing SfC Sites.* University of Arizona.

Bernard, T., Delarue, J., & Naudet, J.-D. (2011). *On 'nailing' what works through impact evaluations: Lessons from experience of AFD.*

Boulton, J. (2012). *Strategy for a Complex World. In J. Verity (Ed.), The New Strategic Landscape: Innovative Perspectives on Strategy*. : Palgrave Macmillan.

Boulton, J., Allen, P., & Bowman, C. (2013). Embracing Complexity: Oxford University Press.

Bouman, F. J. A. (1995). *Rotating and Accumulating Savings and Credit Associations: A Development Perspective.* World Development, 23(3), 371-384.

Checkland, P. (1999). *Soft Systems Methodology in Action* London: Wiley.

Coady Institute. (2012). *A Study of the Meso-level Impacts of Savings Groups* Nairobi: FSD Kenya.

DAI. (2010). *Group Savings and Loan Associations: Impact Study*. Nairobi: FSD Kenya.

Duvendack, M., Palmer-Jones, R., Copestake, J. G., Hooper, L., Loke, Y., & Rao, N. (2011). *What is the evidence of the impact of microfinance on the well-being of poor people?* . London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Elliott, H. (2012). *Ethnographic Study of Meso-Leve Impacts of Savings Groups in Malanga, Coast Province* Unpublished.

FSD Kenya. (2012). *Results of A Study of Post-Project Replication of Groups in COSALO.* Nairobi: FSD Kenya / FSD Updates.

Gasper, D., & Bell, S. (2000). *Evaluating the 'logical framework approach'-towards learning-oriented development evaluation.* Public Administration and Development, 20(1), 17-32.

Georgakopoulou, A. (2010). *Narrative Analysis The Sage Handbook of Sociolinguistics:* Sage.

Heyer, A., & Mas, I. (2011). *Fertile grounds for mobile money: Towards a framework for analysing enabling environments. Enterprise Development and Microfinance, 22(1), 30-44. doi: 10.3362/1755-1986.2011.005*

Hughes, K., & Hutchings, C. (2011). *Can we obtain the required rigour without randomisation? Oxfam GB's non-experimental Global Performance Framework. www.3ieimpact.org: 3ie Working Paper 13.*

Johnson, S. (2004). *Gender norms in financial markets: evidence from Kenya. World Development, 32(8), 1355-1374.*

Johnson, S., Brown, G. K., & Fouillet, C. (2012). *The Search for Inclusion in Kenya's Financial Landscape: The Rift Revealed.* Nairobi: Financial Sector Deepening Trust, Kenya.

Johnson, S., & Sharma, N. (2007). *'Institutionalizing suspicion': The management and governance challenge in user-owned microfinance groups. In T. Dichter & M. Harper (Eds.), What's wrong with microfinance.* Rugby: Intermediate Technology Publications Ltd.

Karlan, D., & Appel, J. (2012). *More Than Good Intentions: Improving the Ways the World's Poor Borrow, Save, Farm, Learn, and Stay Healthy.* New York: Plume Books.

Karlan, D., Goldberg, N., & Copestake, J. (2009). *'Randomized control trials are the best way to measure impact of microfinance programmes and improve microfinance product designs'. Enterprise development and microfinance, 20(3), 167-176.*

Lewis, D., & Mosse, D. (2006). *Brokers and Translators: An Ethnography of Aid and Agencies Bloomfield,* USA: Kumarian Press

Lipsky, M. (2010). *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services.* New York: Russell Sage Foundation

Luhman, J. T., & Boje, D. M. (2001). *What is complexity science? A possible answer from narrative research.* Emergence, 3(1), 158-168.

Malkamaki, M. (2013). *Rule following and financial performance of COSALO saving groups in Nyanza,* Kenya: Unpublished report for FSD Kenya.

Mayne, J. (2008). *Contribution Analysis: An approach to exploring cause and effect http://www.cgiar-ilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf: ILAC.*

Mayne, J. (2011). *Contribution Analysis: Addressing Cause and Effect*. In K. Forss, M. Marra & R. Schwartz (Eds.), *Evaluating the Complex: Attribution, Contribution and Beyond*. New Brunswick (USA) and London (UK): Transaction Publishers.

Mosse, D. (2005). *Cultivating development: an ethnography of aid policy and practice.* London: Pluto Press.

Mule, N., Johnson, S., Hickson, R., & Mwangi, W. (2002). *The Managed ASCA Model: Innovation in Kenya's Microfinance Industry.* Nairobi: Microsave.

Odell, M., & Rippey, P. (2010). *The Permanence and Value of Savings Groups in CARE Kenya's COSAMO Program Nyanza Province,* Kenya: Unpublished. .

Odell, M. L., & Rippey, P. (2011). *Beyond Financial Services: The Permanence and Value of Savings Groups in CARE Kenya's COSAMO Programme.* Geneva: Aga Khan Foundation.

Pande, R., Cole, S., Sivasankaran, A., Bastian, G., & Durlacher, K. (2012). *Does poor people's access to formal banking services raise their incomes? London: EPPI-Centre, Social Science Research Unit, Institute of Education,* University of London.

Patton, M. Q. (1994). *Developmental Evaluation. Evaluation Practice, 15(3), 311-320.*

Snowden, D. (2002). *Complex acts of knowing: paradox and descriptive self-awareness. Journal of Knowledge Management, 6(2), 100-111.*

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the Range of Designs and Methods for Impact Evaluations. London: DFID Working Paper 38.*

Stone, R., Johnson, S., & Hayes, J. (2010). *Financial Sector Deepening Kenya Impact Assessment* Nairobi: FSD Kenya.

Systems Concepts in Evaluation; *An Expert Anthology. (2006).* American Evaluation Association.

Vogel, I. (2012). *Review of the use of 'Theory of Change' in international development* DFID London.

von Bertalanffy, L. (1969). *General System Theory:* George Brazillier.

White, H., & Phillips, D. (2012). A*ddressing attribution of cause and effect in small n impact evaluations: towards an integrated framework. www.3ieimpact.org: 3ie Working Paper 15.*

Woolcock, M. (2009). T*oward a plurality of methods in project evaluation: a contextualised approach to understanding impact trajectories and efficacy. Journal of Development Effectiveness, 1(1), 1 - 14.*

Notes

Notes

FSD Kenya
Financial Sector Deepening

info@fsdkenya.org • www.fsdkenya.org
FSD Kenya is an independent Trust established to support the development of inclusive financial markets in Kenya
5th floor, KMA Centre • Junction of Chyulu Road and Mara Road, Upper Hill • PO Box 11353, 00100 Nairobi, Kenya
T +254 (20) 2923000 • **C** +254 (724) 319706, (735) 319706